

# An extension of the Dirac and Gotay-Nester theories of constraints for Dirac dynamical systems

Hernán Cendra<sup>a,1</sup>, María Etchehoury<sup>b</sup>  
and Sebastián J. Ferraro<sup>c</sup>

*a* Departamento de Matemática  
Universidad Nacional del Sur, Av. Alem 1253  
8000 Bahía Blanca and CONICET, Argentina.  
uscendra@criba.edu.ar

*b* Laboratorio de Electrónica Industrial, Control e Instrumentación,  
Facultad de Ingeniería, Universidad Nacional de La Plata and  
Departamento de Matemática, Facultad de Ciencias Exactas, Universidad  
Nacional de La Plata.  
CC 172, 1900 La Plata, Argentina.  
marila@mate.unlp.edu.ar

*c* Departamento de Matemática and Instituto de Matemática Bahía Blanca  
Universidad Nacional del Sur, Av. Alem 1253  
8000 Bahía Blanca,  
and CONICET, Argentina  
sferraro@uns.edu.ar

## Abstract

This paper extends the Gotay-Nester and the Dirac theories of constrained systems in order to deal with Dirac dynamical systems in the integrable case. Integrable Dirac dynamical systems are viewed as constrained systems where the constraint submanifolds are foliated, the case considered in Gotay-Nester theory being the particular case where the foliation has only one leaf. A Constraint Algorithm for Dirac dynamical systems (CAD), which extends the Gotay-Nester algorithm, is developed. Evolution equations are written using a Dirac bracket adapted to the foliations and an abridged total energy which coincides with the total Hamiltonian in the particular case considered by Dirac. The interesting example of LC circuits is developed in detail. The paper emphasizes the point of view that Dirac and Gotay-Nester theories are dual and that using a combination of results from both theories may have advantages in dealing with a given example, rather than using systematically one or the other.

---

<sup>1</sup>Corresponding author. Fax: +54-221-424-5875

# 1 Introduction

The main purpose of this paper is to study an extension of the Dirac theory of constraints (Dirac [1950, 1958, 1964]) and the Gotay-Nester theory (Gotay, Nester, and Hinds [1978]; Gotay and Nester [1979]) for the case in which the equation of motion is a Dirac dynamical system on a manifold  $M$ ,

$$(x, \dot{x}) \oplus d\mathcal{E}(x) \in D_x, \quad (1.1)$$

where  $D \subseteq TM \oplus T^*M$  is a given *integrable* Dirac structure on  $M$ , which gives a foliation of  $M$  by presymplectic leaves, and  $\mathcal{E}$  is a given function on  $M$ , called the *energy*. One of the features of systems like (1.1) is that they can represent some constrained systems where the constraints appear even if the Lagrangian is nonsingular, as it happens in the case of integrable nonholonomic systems. We will work under explicit hypotheses about regularity, locality and integrability, although some of our results can be applied in more general cases, as indicated in the paragraph *Hypotheses*, below in this section.

In the Gotay-Nester theory the starting point is an equation of the form

$$\omega(x)(\dot{x}, \cdot) = d\mathcal{E}(x) \quad (1.2)$$

where  $\omega$  is a presymplectic form on  $M$ . In fact, in the Gotay-Nester theory the more general case where  $d\mathcal{E}(x)$  is replaced by an arbitrary closed 1-form  $\alpha$  is considered, but this will not be relevant here.

**Example: Euler-Lagrange and Hamilton's equations.** Let  $L: TQ \rightarrow \mathbb{R}$  be a Lagrangian, degenerate or not. Let  $\mathcal{E}(q, v, p) = pv - L(q, v)$  and let  $\omega \in \Omega^2(TQ \oplus T^*Q)$  be the presymplectic form  $\omega = dq^i \wedge dp_i$  on the Pontryagin bundle  $M = TQ \oplus T^*Q$ . Then the Euler-Lagrange equations are written equivalently in the form of equation (1.2) with  $x = (q, v, p)$ . In fact, we have, in local coordinates,

$$i_{(\dot{q}, \dot{v}, \dot{p})} dq^i \wedge dp_i = \dot{q}^i dp_i - \dot{p}_i dq^i \quad (1.3)$$

$$d\mathcal{E} = \frac{\partial \mathcal{E}}{\partial q^i} dq^i + \frac{\partial \mathcal{E}}{\partial p_i} dp_i + \frac{\partial \mathcal{E}}{\partial v^i} dv^i \quad (1.4)$$

$$= -\frac{\partial L}{\partial q^i}(q, v) dq^i + v^i dp_i + \left( p_i - \frac{\partial L}{\partial v^i} \right) dv^i \quad (1.5)$$

Using equations (1.3)–(1.5) we can easily see that (1.2) becomes

$$\begin{aligned} \dot{q}^i &= v^i \\ \dot{p}_i &= \frac{\partial L}{\partial q^i}(q, v) \\ 0 &= p_i - \frac{\partial L}{\partial v^i}(q, v), \end{aligned}$$

which is clearly equivalent to the Euler-Lagrange equations. The case  $\mathcal{E}(q, v, p) = H(q, p)$ , where  $H$  is a given Hamiltonian on  $T^*Q$  of course gives Hamilton's equations.

The idea of using the Pontryagin bundle to write such fundamental equations of physics appears in [Livens \[1919\]](#); [Skinner \[1983\]](#); [Skinner and Rusk \[1983a,b\]](#); [Cendra, Marsden, Pekarsky, and Ratiu \[2003b\]](#); [Yoshimura and Marsden \[2006a,b\]](#).

Equation (1.2) is equivalent to equation (1.1) in the case in which the Dirac structure on  $M$  is the one naturally associated to  $\omega$ , denoted  $D_\omega$ , in which case the foliation of  $M$  has obviously only one presymplectic leaf. In general, equation (1.1) may be considered as a collection of equations of the type (1.2), one for each presymplectic leaf of the Dirac structure. However, in order to study the initial condition problem this approach might not be appropriate, because different initial conditions might belong to different presymplectic leaves and therefore correspond to different equations of the type (1.2), which is not the usual situation.

The algorithm of the Gotay-Nester theory generates a finite sequence of secondary constraint submanifolds  $M \supseteq M_1 \supseteq \dots \supseteq M_c$ . The final constraint submanifold  $M_c$  has the property that every solution curve of (1.2) is contained in  $M_c$  (in fact, it is the smallest submanifold having that property). Equations of motion are given by restricting the variable  $(x, \dot{x})$  in equation (1.2) to  $TM_c$ , and existence of solution curves  $x(t)$  for a given initial condition  $x_0 \in M_c$  is guaranteed, under the condition that the kernel of

$$\omega^\flat(x)|_{T_x M_c}: T_x M_c \rightarrow T_x^* M$$

has locally constant dimension. An important point in the Gotay-Nester approach is that the equations defining the secondary constraint submanifolds are written in terms of  $\omega$  (see formula (2.5)), which makes the whole algorithm invariant under a group of transformations preserving  $\omega$ .

In order to solve (1.1), we will develop a constraint algorithm that we call CAD (*constraint algorithm for Dirac dynamical systems*). The CAD extends the Gotay-Nester algorithm and gives a sequence of secondary foliated constraint submanifolds. The secondary foliated constraint submanifolds can be written in terms of the Dirac structure using formula (5.6), which generalizes (2.5).

Let  $M \subseteq N$  be an embedding of  $M$  into a symplectic manifold  $(N, \Omega)$ , in such a way that the presymplectic leaves of  $M$  are presymplectic submanifolds of  $N$  and let, by slight abuse of notation,  $\mathcal{E}: N \rightarrow \mathbb{R}$  denote an arbitrary extension of the given energy on  $M$ . This kind of framework can be naturally constructed in many examples. Then one can consider, at least locally, that  $M$  is a primary foliated constraint submanifold defined by equating certain constraints to 0, which generalizes the situation of the Dirac theory of constraints. One of the results of this paper establishes that there exists a Dirac bracket whose symplectic leaves are *adapted* to the foliations of the primary and final foliated constraint submanifolds. We also prove that the equations of motion on the final foliated constraint submanifold can be nicely written in Hamiltonian form with respect to this Dirac bracket and the *abridged total energy*, which we introduce in section 3.4.

Our extension of the Gotay-Nester and Dirac theories has especially simple and interesting features when applied to LC circuits. We show that the algorithm stops either at  $M_1$  or  $M_3$ , and concrete physically meaningful formulas for the first three

constraint submanifolds and for the evolution equations can be given for this case. Other geometric treatments of LC circuits can be found, for instance, in Bloch and Crouch [1999]; Yoshimura and Marsden [2006a,b]; Cendra, Marsden, Ratiu, and Yoshimura [2011].

Systems of the type (1.1) are important also in the not necessarily integrable case since they represent, for instance, the Lagrange-d'Alembert equations of non-holonomic mechanics. More references will be given in section 4.1. We remark that even though the main focus of this paper is the case where the Dirac structure is integrable, many results, most notably the constraint algorithm CAD and the treatment of nonholonomic systems in section 5, are proven for a not necessarily integrable Dirac structure.

Gotay-Nester's theory is more geometric than Dirac's and provides a certain notion of duality or correspondence between the two, in the spirit of the dualities between submanifolds and the equations defining them, manifolds and their ring of functions, Hamiltonian vector fields and their Hamiltonian or Poisson brackets and the collection of their symplectic leaves. Emphasizing this point of view in a common geometric language and showing that a combination of results from both theories may have advantages in a given example, rather than using one or the other, is an aspect of the paper.

From this point of view, the notion of second class constraints is related to the notion of second class constraint submanifold, the latter being determined in this paper as a submanifold that is tangent to a second class subbundle of a certain tangent bundle. They are symplectic submanifolds (Śniatycki [1974]; Bursztyn [2010]). The first class constraint submanifolds are coisotropic submanifolds (Bursztyn [2010]).

The presence of primary and secondary first class constraints, which in Dirac's work is connected to the principle of determinacy and to the notion of physical variables, is explicitly related here to certain manifolds  $\tilde{M}$  and  $\bar{M}$  in a commutative diagram, in Theorem 3.19. The diagonal arrow in this diagram, which is a submersion, is dual to the diagonal arrow in the diagram in Lemma 3.17, which is a surjective Poisson algebra homomorphism. In some examples, instead of applying the Dirac method to write equations of motion, it may be simpler to proceed in two stages. First apply the Gotay-Nester algorithm to find the final constraint submanifold, which, as proven in the Gotay-Nester theory, coincides with the submanifold defined by equating the final constraints obtained in the Dirac algorithm to 0. Then switch to the Dirac viewpoint and write equations of motion using the Dirac bracket. Also, in order to calculate the Dirac bracket at points of the final constraint submanifold, using the symplectic leaf that contains that submanifold may be easier than using Dirac's expression.

**The point of view of IDEs.** In order to study the problem of finding solution curves of general Dirac dynamical systems (1.1) the more general theory of *Implicit Differential Equations* may be useful. An Implicit Differential Equation (IDE) on a manifold  $M$  is written as

$$\varphi(x, \dot{x}) = 0. \quad (1.6)$$

A solution of (1.6) at a point  $x$  is a vector  $(x, v) \in T_x M$  satisfying (1.6). A *solution curve*, say  $x(t)$ ,  $t \in (a, b)$ , must satisfy, by definition, that  $(x(t), \dot{x}(t))$  is a solution at  $x(t)$  for all  $t \in (a, b)$ .

Basic questions such as existence, uniqueness or extension of solutions for a given initial condition are not completely answered yet, although many partial results have been established for certain classes of IDE. One approach is to use a *constraint algorithm*, which consists of a decreasing sequence of constraint submanifolds  $M \supseteq M_1 \supseteq \dots \supseteq M_c$  defined as follows,

$$M_{k+1} := \{x \in M_k \mid \text{there exists } (x, v) \in TM_k \text{ such that } \varphi(x, v) = 0\},$$

with  $M_0 := M$ . This algorithm, which obviously uses only the differentiable structure regardless of any other structure which may be present, like presymplectic, Poisson or Dirac structures, represents a differential geometric aspect underlying the algorithms of Gotay-Nester, Dirac or CAD. To ensure that each  $M_k$  is a submanifold and that the algorithm stops after a finite number of steps one may choose to assume certain conditions of the type “locally constant rank” conditions. Then the original IDE is reduced to an equivalent ODE depending on parameters on the *final constraint submanifold*  $M_c$ . In fact, by construction,  $M_c$  is characterized by the property that it is the smallest submanifold that contains all solutions curves of the given IDE. Therefore, if  $M_c$  is empty, there are no solution curves.

Dirac’s original work has a wide perspective from the physical point of view, with connections to classical, quantum and relativistic mechanics. However, from the point of view of abstract IDE theory, and in a very concise way, we may say that a combination of the Dirac and the Gotay-Nester methods shows how to transform a given Gotay-Nester equation (1.2) into an equivalent ODE depending on parameters (3.24) on a final constraint submanifold, while what we show in this paper is how to transform a given Dirac dynamical system (1.1) into an equivalent ODE (7.11) depending on parameters on a final foliated constraint submanifold.

Some more comments on the connection between Dirac’s and Gotay-Nester ideas and IDE are in order. We can compare Dirac [1950]; Gotay, Nester, and Hinds [1978]; Gotay and Nester [1979]; Rabier and Rheinboldt [1994]; Pritchard [2003]; Cendra and Etchechoury [2006], to see how the idea of applying a constraint algorithm works in different contexts. In Cendra and Etchechoury [2006], one works in the realm of subanalytic sets; in Gotay, Nester, and Hinds [1978] and Gotay and Nester [1979] one works with presymplectic manifolds; in Pritchard [2003] one works with complex algebraic manifolds; Dirac [1950] uses Poisson brackets; in Rabier and Rheinboldt [1994] some degree of differentiability of the basic data is assumed, and, besides, some constant rank hypotheses are added, essentially to ensure applicability of a certain constant rank theorem. Some relevant references for general IDEs connected to physics or control theory, which show a diversity of geometric or analytic methods or a combination of both are Cariñena and Rañada [1984]; Gràcia and Pons [1991, 1992]; Mendella, Marmo, and Tulczyjew [1995]; de León and de Diego [1996]; Ibort, de León, Marmo, and Martín de Diego [1996]; Marmo, Mendella, and Tulczyjew [1997]; Delgado-Téllez and Ibort [2003].

**Hypotheses.** As for regularity, we will work in the  $C^\infty$  category. Throughout the paper we assume that the subsets  $M_k$  appearing in sequences of the type  $M \supseteq M_1 \supseteq \cdots \supseteq M_c$  that are generated by some constraint algorithm, are submanifolds regularly defined by equating some functions (constraints) to 0, and that the sequence stops. More regularity conditions like Assumptions 3.8, 3.18, 3.26, 3.27, 7.1, and some others will be introduced when needed along the paper.

Our results will be of a local character, but for some of them, like the notion of second class submanifolds, it is indicated how to define them globally. The usage of local coordinates is almost entirely avoided and basic facts in symplectic geometry or Poisson algebra arguments are used instead.

The condition of integrability of the Dirac structure appearing in (1.1) has its own interest. However, the CAD does not assume integrability and can be applied for instance to general nonholonomic systems. On the other hand, in the non-integrable case certain brackets that we study would not satisfy the Jacobi identity, and will not be studied in this paper.

**Structure of the paper.** The first part of the paper, which includes sections 2 and 3, contains a review of the Dirac and Gotay-Nester methods.

In sections 4 to 7 we develop our main results, extending the Gotay-Nester and Dirac theories. Section 4 is devoted to a review of basic facts on Dirac structures and Dirac dynamical systems (Cendra, Marsden, Ratiu, and Yoshimura [2011]). The notion of Dirac structure (Courant and Weinstein [1988]; Courant [1990]; Bursztyn and Crainic [2009]) gives a new possibility of understanding and extending the theory of constraints, which is the main purpose of the present paper. In section 5 we develop our CAD algorithm. We do it for the general case of a not necessarily integrable Dirac structure so that one can apply CAD to general nonholonomic systems. In section 6 we study the example of integrable nonholonomic systems, and LC circuits are viewed as a particular case. In section 7 we show how to extend the Dirac theory for the case of a Dirac dynamical system (1.1).

## 2 A brief review of the Dirac and Gotay-Nester theories

In this section we review, without proof, basic facts belonging to the Dirac and the Gotay-Nester theories.

The Dirac method starts with a given submanifold  $M$ , called the primary constraint submanifold, of a symplectic manifold  $(N, \Omega)$  defined by equating the primary constraints to 0, and a given energy  $\mathcal{E}: N \rightarrow \mathbb{R}$ . In the original work of Dirac  $N = T^*Q$ ,  $\Omega$  is the canonical symplectic form,  $M$  is the image of the Legendre transformation and  $\mathcal{E}$  is the Hamiltonian defined by Dirac. However, locally, any symplectic manifold is an open subset of a cotangent bundle, therefore the original Dirac formalism can be applied to the, seemingly more general, case described above, at least for the local theory.

An interesting variant of the original situation considered by Dirac is the following. Consider the canonical symplectic manifold  $N = T^*TQ$  with the canonical

symplectic form  $\Omega$ , and let the primary constraint be  $M = TQ \oplus T^*Q$ , canonically embedded in  $N$  via the map given in local coordinates  $(q, v, p, \nu)$  of  $N$  by  $\varphi(q, v, p) = (q, v, p, 0)$ . In particular,  $M$  is defined regularly by the equation  $\nu = 0$ . If  $\omega$  is the presymplectic form on  $M$  obtained by pulling back the canonical symplectic form of  $T^*Q$  using the natural projection, then one can show that  $\varphi^*\Omega = \omega$ . The embedding  $\varphi$  is globally defined (see Appendix A for details). The number  $pv$  is a well-defined function on  $M$  and it can be naturally extended to a function on a chart with coordinates  $(q, v, p, \nu)$ , but this does not define a global function on  $N$  consistently. On the other hand, it can be extended to a smooth function on  $N$  using partitions of unity and any such extension will give consistent equations of motion. In this paper we will not consider global aspects. In any case, for a given Lagrangian  $L: TQ \rightarrow \mathbb{R}$  we can take  $\mathcal{E} = pv - L(q, v)$ .

*Remark 2.1.* For a given presymplectic manifold  $(M, \omega)$  one can always find an embedding  $\varphi$  into a symplectic manifold  $(P, \Omega)$  such that  $\varphi^*\Omega = \omega$ . Moreover, this embedding can also be chosen such that it is coisotropic, meaning that  $\varphi(M)$  is a coisotropic submanifold of  $P$  (see Gotay and Śniatycki [1981/82]). However, we should mention that the embedding given above is not coisotropic.

The Dirac and the Gotay-Nester algorithms can be studied independently. On the other hand, they are related as follows. For a given system (1.2) choose a symplectic manifold  $(N, \Omega)$  in such a way that  $(M, \omega)$  is a presymplectic submanifold and  $\mathcal{E}$  (using a slight abuse of notation) is an arbitrary extension of  $\mathcal{E}: M \rightarrow \mathbb{R}$  to  $N$ . Moreover, assume that  $M$  is defined regularly by a finite set of equations  $\phi_i^{(0)} = 0$ ,  $i = 1, \dots, a_0$ , where each  $\phi_i^{(0)}$  is a *primary constraint*. The Dirac algorithm gives a sequence of secondary constraints  $\phi_i^{(k)}$ ,  $i = 1, \dots, a_k$ , which defines regularly a sequence of secondary constraint submanifolds  $M_k$ ,  $k = 1, \dots, c$ , by equations  $\phi_i^{(k)} = 0$ ,  $i = 1, \dots, a_k$ , which coincide with the ones given in the Gotay-Nester algorithm.

## 2.1 A brief review of Dirac's theory

Dirac's theory of constraints has been extensively studied from many different points of view and extended in several directions. Part of those developments in the spirit of geometric mechanics is contained in the following references, but the list is far from being complete, Sudarshan and Mukunda [1974]; Gotay, Nester, and Hinds [1978]; Gotay and Nester [1979, 1980]; Mukunda [1980, 1987]; Skinner [1983]; Skinner and Rusk [1983a,b]; Cariñena and Rañada [1984, 1993, 1995]; Cariñena, Gomis, Ibort, and Román [1985]; Cantrijn, Cariñena, Crampin, and Ibort [1986]; van der Schaft [1987]; Cariñena, López, and Román-Roy [1988]; Cariñena [1990]; Henneaux and Teitelboim [1992]; Krupková [1994]; de León, de Diego, and Pitanga [1995]; Krupková [1996]; Marmo, Mendella, and Tulczyjew [1997]; Ibort, de León, Marrero, and Martín de Diego [1999]; Barbero-Liñán and Muñoz-Lecanda [2009]; Crampin and Mestdag [2011a,b].

As we have explained above we will work in a general context of a given symplectic manifold  $(N, \Omega)$ ,  $M \subseteq N$  and  $\mathcal{E}: N \rightarrow \mathbb{R}$ , where the primary constraint



submanifold  $M$  is regularly defined by equations  $\phi_i^{(0)} = 0$ ,  $i = 1, \dots, a_0$  on  $N$ .

The Dirac constraint algorithm goes as follows. One defines the *total energy*  $\mathcal{E}_T = \mathcal{E} + \lambda_{(0)}^i \phi_i^{(0)}$ . The preservation of the primary constraints is written  $\{\phi_i^{(0)}, \mathcal{E}_T\}(x) = 0$ ,  $i = 1, \dots, a_0$ ,  $x \in M_0$ , or

$$\{\phi_i^{(0)}, \mathcal{E}\}(x) + \lambda_{(0)}^j \{\phi_i^{(0)}, \phi_j^{(0)}\}(x) = 0, \quad i, j = 1, \dots, a_0, \quad x \in M_0.$$

Then  $M_1$  is defined by the condition that  $x \in M_1$  if and only if there exists  $\lambda_{(0)} = (\lambda_{(0)}^1, \dots, \lambda_{(0)}^{a_0})$  such that the system of equations  $\phi_i^{(0)}(x) = 0$ ,  $\{\phi_i^{(0)}, \mathcal{E}_T\}(x) = 0$ ,  $i = 1, \dots, a_0$ , is satisfied. The submanifold  $M_1$  is defined by equations  $\phi_i^{(1)} = 0$ ,  $i = 1, \dots, a_1$ , where each  $\phi_i^{(1)}$  is a *secondary constraint*, by definition. By proceeding iteratively one obtains a sequence  $M_0 \supseteq M_1 \supseteq \dots \supseteq M_c$ . Then there are *final constraints*, say  $\phi_i^{(c)}$ ,  $i = 1, \dots, a_c$ , defining a submanifold  $M_c$  by equations  $\phi_i^{(c)} = 0$ ,  $i = 1, \dots, a_c$ , called the *final constraint submanifold*, and the following condition is satisfied: for each  $x \in M_c$  there exists  $(\lambda_{(0)}^1, \dots, \lambda_{(0)}^{a_0})$  such that

$$\{\phi_i^{(c)}, \mathcal{E}\}(x) + \lambda_{(0)}^j \{\phi_i^{(c)}, \phi_j^{(0)}\}(x) = 0, \quad i = 1, \dots, a_c, \quad j = 1, \dots, a_0. \quad (2.1)$$

For each  $x \in M_c$  the space of solutions of the linear system of equations (2.1) in the unknowns  $\lambda_{(0)}^j$  is an affine subspace of  $\mathbb{R}^{a_0}$ , called  $\Lambda_x^{(c)}$ , whose dimension is a locally constant function  $d^{(c)}(x) = a_0 - \text{rank}(\{\phi_i^{(c)}, \phi_j^{(0)}\}(x))$ . One can locally choose  $d^{(c)}(x)$  unknowns as being free parameters and the rest will depend affinely on them. Then the solutions of (2.1) form an affine bundle  $\Lambda^{(c)}$  over  $M_c$ . After replacing  $\lambda_{(0)} \in \Lambda^{(c)}$  in the expression of the total energy, the corresponding Hamiltonian vector field,

$$X_{\mathcal{E}_T}(x) = X_{\mathcal{E}}(x) + \lambda_{(0)}^j X_{\phi_j^{(0)}}(x), \quad (2.2)$$

$x \in M_c$ , which will depend on the free unknowns, will be tangent to  $M_c$ . Its integral curves, for an arbitrary choice of a time dependence of the free unknowns, will be solutions of the equations of motion, which is the main property of the final constraint submanifold  $M_c$  from the point of view of classical mechanics. The lack of uniqueness of solution for a given initial condition in  $M_c$ , given by the presence of free parameters, indicates, according to Dirac, the nonphysical character of some of the variables. In our context the physical variables will be given a geometric meaning.

*Remark 2.2.* Dirac introduces the notion of weak equality for functions on  $T^*Q$ . Two such functions are *weakly equal*, denoted  $f \approx g$ , if  $f|_{M_c} = g|_{M_c}$ . Then, for instance  $\phi_j^{(k)} \approx 0$ . If  $f \approx 0$  then  $f = \nu^i \phi_i^{(c)}$ , for some functions  $\nu^i$  on  $T^*Q$  and conversely. Since we use the notion of a constraint submanifold, in particular the final constraint submanifold, we prefer not to use the notation  $\approx$ .

Now let us make some comments on the notions of first class and second class constraints. The rank of the skew-symmetric matrix  $(\{\phi_i^{(c)}, \phi_j^{(c)}\}(x))$ ,  $i, j = 1, \dots, a_c$ , is necessarily even, say,  $2s$ , and it is assumed to be constant on  $M_c$ , as part of



our regularity conditions; for our results we will assume later a stronger condition (Assumption 3.27). One can choose, among the  $\phi_i^{(c)}$ ,  $i = 1, \dots, a_c$ ,  $2s$  functions  $\chi_j$ ,  $j = 1, \dots, 2s$ , and replace the rest of the  $\phi_i^{(c)}$  by  $\phi_i^{(c)} + \alpha_i^j \chi_j$  with appropriate functions  $\alpha_i^j$  in such a way that  $\psi_i = 0$ ,  $\chi_j = 0$ , define  $M_c$  regularly and, besides,  $\{\psi_i, \psi_{i'}\}(x) = 0$ ,  $\{\psi_i, \chi_j\}(x) = 0$ ,  $\det(\{\chi_j, \chi_{j'}\}(x)) \neq 0$ , for  $i, i' = 1, \dots, a_c - 2s$ ,  $j, j' = 1, \dots, 2s$  and  $x \in M_c$ . The  $\phi_j^{(c)}$  are linear combinations with smooth coefficients of the  $\chi_j$  and  $\psi_i$ , and conversely. The functions  $\chi_j$ ,  $j = 1, \dots, 2s$ , are called *second class constraints* and the functions  $\psi_i$ ,  $i = 1, \dots, a_c - 2s$ , are called *first class constraints*.

More generally, any function  $\rho$  on  $T^*Q$  satisfying  $\rho|_{M_c} = 0$ ,  $\{\rho, \psi_i\}|_{M_c} = 0$ ,  $\{\rho, \chi_j\}|_{M_c} = 0$ , is a first class constraint with respect to the submanifold  $M_c$ , by definition. Any function  $g$  on  $T^*Q$  satisfying  $\{g, \psi_i\}|_{M_c} = 0$ ,  $\{g, \chi_j\}|_{M_c} = 0$ , is a *first class function*, by definition. For instance, the total energy  $\mathcal{E}_T$  is a first class function.

Now define the energy  $\mathcal{E}_c$  in terms of  $\psi_i$ ,  $\chi_j$ ,  $i = 1, \dots, a_c - 2s$ ,  $j = 1, \dots, 2s$ , as

$$\mathcal{E}_c = \mathcal{E} + \lambda^i \psi_i + \mu^j \chi_j.$$

The preservation of the constraints for the evolution generated by  $\mathcal{E}_c$  can be rewritten as  $\{\psi_i, \mathcal{E}_c\}(x) = 0$ , which is equivalent to  $\{\psi_i, \mathcal{E}\}(x) = 0$  for all  $x \in M_c$ , and  $\{\chi_j, \mathcal{E}_c\}(x) = 0$ , for all  $x \in M_c$ . The latter is equivalent to

$$\{\chi_i, \mathcal{E}\}(x) + \mu^j \{\chi_i, \chi_j\}(x) = 0, \quad i, j = 1, \dots, 2s,$$

for all  $x \in M_c$ , which determines the  $\mu^j$  as well-defined functions on  $M_c$ . Then the solutions  $(\mu(x), \lambda)$  form an affine bundle with base  $M_c$  and whose fiber, parametrized by the free parameters  $\lambda$ , has dimension  $a_c - 2s$ .

Any section  $(\mu(x), \lambda(x))$  of this bundle determines  $\mathcal{E}_c$  as a first class function. This means that  $X_{\mathcal{E}_c}(x) \in T_x M_c$ , for each  $x \in M_c$ , and therefore a solution curve of  $X_{\mathcal{E}_c}$  is contained in  $M_c$  provided that the initial condition belongs to  $M_c$ . The function  $\mathcal{E}_c$  is essentially the *extended Hamiltonian* defined by Dirac.

Dirac defines an interesting bracket, now called the *Dirac bracket*,

$$\{F, G\}^* = \{F, G\} - \{F, \chi_i\} c^{ij} \{\chi_j, G\},$$

which is defined on an open set in  $T^*Q$  containing  $M_c$ , where  $c^{ij}$ , which by definition is the inverse matrix of  $\{\chi_i, \chi_j\}$ , is defined. The Dirac bracket is a Poisson bracket and has the important property that for *any* function  $F$  on  $T^*Q$ , the condition  $\{F, \chi_j\}^* = 0$ ,  $j = 1, \dots, 2s$ , is satisfied on a neighborhood of  $M_c$ , which implies that  $\dot{F} = \{F, \mathcal{E}_c\} = \{F, \mathcal{E}_c\}^*$ , for any function  $F$ . Besides,  $\{\psi_j, \psi_i\}^* = 0$ ,  $i, j = 1, \dots, a_c - 2s$ , on  $M_c$ . Because of this, one may say that, with respect to the Dirac bracket, all the constraints  $\chi_j$ ,  $j = 1, \dots, 2s$  and  $\psi_i$ ,  $i = 1, \dots, a_c - 2s$ , behave like first class constraints, with respect to  $M_c$ .

## 2.2 A brief review of the Gotay-Nester algorithm

We will need just a few basic facts from the Gotay-Nester theory.

In order to find solution curves to (1.2) we can apply the general algorithm for IDEs described in the introduction, and the final IDE can be written

$$\begin{aligned}\omega(x)(\dot{x}, \cdot) &= d\mathcal{E}(x) \\ (x, \dot{x}) &\in TM_c.\end{aligned}$$

Let  $\omega_c$  be the pullback of  $\omega$  to  $M_c$ . If  $\omega_c$  is symplectic, one obtains the simpler equivalent equation

$$\omega_c(x)(\dot{x}, \cdot) = d(\mathcal{E}|_{M_c}) \quad (2.3)$$

on  $M_c$ , which is in Hamiltonian form. However, one must be aware that if  $\omega_c$  is degenerate then (2.3) is not equivalent in general to (2.2).

The equation on  $M_c$ ,

$$\omega(x)(X, \cdot) = d\mathcal{E}(x), \quad (2.4)$$

where  $X \in TM_c$ , defines an affine distribution on  $M_c$ , more precisely, one has an affine bundle  $S^{(c)}$  with base  $M_c$  whose fiber  $S_x^{(c)}$  at a given point  $x \in M_c$  is, by definition,

$$S_x^{(c)} = \{X \in T_x M_c \mid (2.4) \text{ is satisfied}\}.$$

The equivalence between the Dirac and the Gotay-Nester algorithms can be made explicit as an isomorphism of affine bundles, as follows. The affine bundles  $S^{(c)} \rightarrow M_c$  and  $\Lambda^{(c)} \rightarrow M_c$  are isomorphic over the identity on  $M_c$ , more precisely, the isomorphism  $\Lambda^{(c)} \rightarrow S^{(c)}$  is given by  $\lambda_{(0)} \rightarrow X_{\mathcal{E}_T}$ , using equation (2.2). In particular, the rank of  $S^{(c)}$  is  $d^{(c)}$ .

**Describing the secondary constraints using  $\omega$ .** The constraint manifolds  $M_k$  defined by the algorithm can be described by *equations written in terms of the presymplectic form  $\omega$* , which is a simple but important idea, because those equations will obviously be invariant under changes of coordinates preserving  $\omega$ . Depending on the nature of  $\omega$  one may obtain analytic, smooth, linear, etc., equations, which may simplify matters in given examples. The context of reflexive Banach manifolds is used in Gotay, Nester, and Hinds [1978] and Gotay and Nester [1979].

The condition defining the subsets  $M_{k+1}$ ,  $k = 0, 1, \dots$  namely,

$$i_v \omega(x) = d\mathcal{E}(x) \text{ has at least one solution } v \in T_x M_k,$$

is equivalent to

$$d\mathcal{E}(x) \in (T_x M_k)^\flat$$

or, since  $(T_x M_k)^\flat = ((T_x M_k)^\omega)^\circ$ ,

$$\langle d\mathcal{E}(x), (T_x M_k)^\omega \rangle = \{0\}. \quad (2.5)$$

### 3 First and second class constraints, constraint submanifolds and equations of motion

In this section we will describe some results on first class and second class constraints and constraint submanifolds and also equations of motion which will be useful for our extension of those notions, to be developed in section 7.

### 3.1 Linear symplectic geometry

The following results about linear symplectic geometry are an essential part of many of the arguments that we use in the paper, since under our strong regularity assumptions many of them are essentially of a linear character.

**Lemma 3.1.** *Let  $(E, \Omega)$  be a symplectic vector space of dimension  $2n$ ,  $V \subseteq E$  a given subspace. For a given basis  $\alpha_i$ ,  $i = 1, \dots, r$  of  $V^\circ$ , let  $X_i = \alpha_i^\sharp$ ,  $i = 1, \dots, r$ . Then the rank of the matrix  $[\alpha_i(X_j)]$  is even, say  $2s$ , and  $X_i$ ,  $i = 1, \dots, r$  form a basis of  $V^\Omega$ . Moreover, the basis  $\alpha_i$ ,  $i = 1, \dots, r$  can be chosen such that for all  $j = 1, \dots, r$*

$$\begin{aligned}\alpha_i(X_j) &= \delta_{i,j-s}, \quad 1 \leq i \leq s \\ \alpha_i(X_j) &= -\delta_{i-s,j}, \quad s+1 \leq i \leq 2s \\ \alpha_i(X_j) &= 0, \quad 2s < i \leq r.\end{aligned}$$

**Proof.** Consider the subspace  $V^\Omega = (V^\circ)^\sharp$ . By a well-known result there is a basis  $X_i$ ,  $i = 1, \dots, r$  of  $V^\Omega$  such that for all  $j = 1, \dots, r$

$$\begin{aligned}\Omega(X_i, X_j) &= \delta_{i,j-s}, \quad 1 \leq i \leq s \\ \Omega(X_i, X_j) &= -\delta_{i-s,j}, \quad s+1 \leq i \leq 2s \\ \Omega(X_i, X_j) &= 0, \quad 2s < i \leq r\end{aligned}$$

then take  $\alpha_i = X_i^\flat$ . The first part of the lemma is easy to prove using this. ■

**Lemma 3.2.** *Let  $\alpha_i$ ,  $i = 1, \dots, r$  be a basis of  $V^\circ$  having the properties stated in Lemma 3.1. Then  $X_i$ ,  $i = 2s+1, \dots, r$  form a basis of  $V \cap V^\Omega$ .*

**Proof.** Let  $X = \lambda^i X_i$  be an arbitrary vector in  $V^\Omega$ . Now  $\lambda^i X_i \in V \cap V^\Omega$  iff  $\alpha_j(X) = \lambda^i(\alpha_j(X_i)) = 0$ ,  $j = 1, \dots, r$ . Since the first  $2s$  columns of the matrix  $[\alpha_i(X_j)]$  are linearly independent and the rest are zero, we must have  $\lambda^i = 0$ , for  $1 \leq i \leq 2s$ , and  $\lambda^i$ ,  $i = 2s+1, \dots, r$  are arbitrary. This means that  $V \cap V^\Omega$  is generated by  $X_i$ ,  $i = 2s+1, \dots, r$ . ■

**Corollary 3.3.**  $\dim V \cap V^\Omega = r - 2s$ .

**Proof.** Immediate from Lemma 3.2. ■

Let  $\omega$  be the pullback of  $\Omega$  to  $V$  via the inclusion.  $\omega = j^*\Omega$ . Then  $(V, \omega)$  is a presymplectic space. In what follows, the  $^\flat$  and  $^\sharp$  operators are taken with respect to  $\Omega$  unless specified otherwise.

**Lemma 3.4.**  $V^\omega = V \cap V^\Omega$ .

**Proof.**  $X \in V^\omega$  iff  $\omega(X, Y) = 0$ ,  $\forall Y \in V$  iff  $\Omega(X, Y) = 0$ ,  $\forall Y \in V$ . This is equivalent to  $X \in V \cap V^\Omega$ . ■

**Lemma 3.5.** *Let  $\gamma_i$ ,  $i = 1, \dots, r$  be a given basis of  $V^\circ$  and let  $Y_i = \gamma_i^\sharp$ ,  $i = 1, \dots, r$ . Let  $\beta \in E^*$  be given. Then the following conditions are equivalent.*

(i)  $\beta(V^\omega) = 0$ .

(ii) The linear system

$$\beta(Y_i) + \lambda^j \gamma_j(Y_i) = 0 \quad (3.1)$$

has solution  $\lambda = (\lambda^1, \dots, \lambda^r)$ .

**Proof.** Let us show that (3.1) has solution  $(\lambda^1, \dots, \lambda^r)$  iff the system

$$\beta(X_k) + \mu^l \alpha_l(X_k) = 0 \quad (3.2)$$

has solution  $(\mu^1, \dots, \mu^r)$ , where  $k, l = 1, \dots, r$  and  $\alpha_l$  is a basis satisfying the conditions of Lemma 3.1. Since  $Y_i, i = 1, \dots, r$  and  $X_k, k = 1, \dots, r$  are both bases of  $V^\Omega$  there is an invertible matrix  $[A_k^i]$  such that  $X_k = A_k^i Y_i$ . Let  $[B_i^l]$  be the inverse of  $[A_k^i]$ , so  $Y_i = B_i^l X_l$ . Assume that (3.1) has solution  $\lambda^j, j = 1, \dots, r$ . We can write (3.1) as

$$\beta(Y_i) + \lambda^j \Omega(Y_j, Y_i) = 0, i = 1, \dots, r.$$

Using this we have that for  $k = 1, \dots, r$

$$\begin{aligned} 0 &= \beta(A_k^i Y_i) + \lambda^j \Omega(Y_j, A_k^i Y_i) = \beta(X_k) + \lambda^j \Omega(Y_j, X_k) \\ &= \beta(X_k) + \lambda^j \Omega(B_j^l X_l, X_k) = \beta(X_k) + \mu^l \Omega(X_l, X_k) \end{aligned}$$

where  $\mu^l = \lambda^j B_j^l$ . This means that the system (3.2) has solution. The converse is analogous. Using this, Lemmas 3.2 and 3.4, and the form of the coefficient matrix  $[\alpha_l(X_k)]$  in Lemma 3.1, the proof that (3.2) has solution  $\mu = (\mu^1, \dots, \mu^r)$  iff  $\beta(V^\omega) = 0$  is easy and is left to the reader. ■

**Lemma 3.6.** Consider the hypotheses in Lemma 3.5. Then the solutions to

$$i_X \omega = \beta|V \quad (3.3)$$

(if any) are precisely  $X = \beta^\sharp + \lambda^j Y_j$ , where  $(\lambda^1, \dots, \lambda^r)$  is a solution to (3.1). A solution to (3.3) exists if and only if  $\beta(V^\omega) = 0$ . If  $\omega$  is symplectic then (3.1) and (3.3) have a unique solution and if, in addition,  $\beta^\sharp \in V$ , then  $\lambda^1 = 0, \dots, \lambda^r = 0$  and  $\beta^\sharp$  coincides with  $X = (\beta|V)^\sharp$  defined by (3.3).

**Proof.** Since  $Y_j, j = 1, \dots, r$  form a basis of  $V^\Omega$  we have that  $(\lambda^1, \dots, \lambda^r)$  is a solution to (3.1) iff  $(\beta + \lambda^j \gamma_j)(V^\Omega) = 0$  iff  $\beta + \lambda^j \gamma_j \in V^\flat$  iff  $\beta^\sharp + \lambda^j Y_j \in V$ . Now, let  $X = \beta^\sharp + \lambda^j Y_j$ , where  $(\lambda^1, \dots, \lambda^r)$  satisfies (3.1). Then we have  $X \in V$  as we have just seen and we also have

$$i_X \omega = (i_X \Omega)|V = X^\flat|V = (\beta + \lambda^j \gamma_j)|V = \beta|V,$$

since  $\gamma_j, j = 1, \dots, r$  generate  $V^\circ$ . We have proven that  $X$  is a solution to (3.3). To prove that every solution  $X$  to (3.3) can be written as before, we can reverse the previous argument. Using this, it is clear that if  $\omega$  is symplectic then (3.1) has unique solution, in particular, we have that  $\det(\gamma_j(Y_i)) \neq 0$ . If, in addition,  $\beta^\sharp \in V$  then  $\lambda^j Y_j = X - \beta^\sharp \in V$ . Since  $Y_j, j = 1, \dots, r$  is a basis of  $V^\Omega$ , using Lemma 3.4 and the fact that  $V^\omega = \{0\}$  we get that  $\lambda^j = 0$  for  $j = 1, \dots, r$ . ■

**Corollary 3.7.** *Let  $\Lambda = \{\lambda \mid \lambda \text{ satisfies (3.1)}\}$ . Then  $\dim \Lambda = r - 2s = \dim \ker \omega$ .*

**Proof.**  $\ker \omega = V^\omega$ , which has dimension  $r - 2s$  from Corollary 3.3 and Lemma 3.4. On the other hand the dimension of the subspace of  $\lambda$  satisfying (3.1) is clearly also  $r - 2s$ , since the coefficient matrix has rank  $2s$ . ■

### 3.2 Poisson-algebraic and geometric study of constraints and constraint submanifolds.

We shall start with the constrained Hamiltonian system  $(N, \Omega, \mathcal{E}, M)$ , where  $(N, \Omega)$  is a symplectic manifold,  $\mathcal{E}: N \rightarrow \mathbb{R}$  is the energy and  $M \subseteq N$  is the primary constraint submanifold. The equation to be solved, according to the Gotay-Nester algorithm, is

$$\omega(x)(X, \cdot) = d\mathcal{E}(x)|_{T_x M}, \quad (3.4)$$

where  $X \in T_x M_c$  and  $x \in M_c$ ,  $M_c$  being the final constraint. Let  $\omega_c$  be the pullback of  $\Omega$  via the inclusion of  $M_c$  in  $N$ . Since  $\omega_c$  is presymplectic,  $\ker \omega_c$  is an involutive distribution. From now on we will assume the following.

**Assumption 3.8.** *The distribution  $\ker \omega_c$  has constant rank and defines a regular foliation  $K_c$ , that is, the natural map  $p_{K_c}: M_c \rightarrow \bar{M}_c$ , where  $\bar{M}_c = M_c/K_c$  is a submersion.*

**Lemma 3.9.** *The following assertions hold:*

- (a) *There is a uniquely defined symplectic form  $\bar{\omega}_c$  on  $\bar{M}_c$  such that  $p_{K_c}^* \bar{\omega}_c = \omega_c$ .*
- (b) *Let  $\bar{X}$  be a given vector field on  $\bar{M}_c$ . Then there is a vector field  $X$  on  $M_c$  that is  $p_{K_c}$ -related to  $\bar{X}$ .*
- (c) *Let  $\bar{f} \in \mathcal{F}(\bar{M}_c)$ . Then there exists a vector field  $X$  on  $M_c$  such that  $X$  is  $p_{K_c}$ -related to  $X_{\bar{f}}$ , and for any such vector field  $X$  the equality  $\omega_c(x)(X, \cdot) = d(p_{K_c}^* \bar{f})(x)$  holds for all  $x \in M_c$ .*
- (d) *Let  $X_{x_0} \in T_{x_0} M_c$ . Then one can choose the function  $\bar{f} \in \mathcal{F}(\bar{M}_c)$  and the vector field  $X$  in (c) in such a way that  $X(x_0) = X_{x_0}$ .*

**Proof.** (a) By definition, the leaves of the foliation  $K_c$  are connected submanifolds of  $M_c$ , that is, each  $p_{K_c}^{-1}(z)$ ,  $z \in \bar{M}_c$ , is connected. For  $z \in \bar{M}_c$ , let  $x \in M_c$  such that  $p_{K_c}(x) = z$ . For  $\bar{A}, \bar{B} \in T_z \bar{M}_c$ , as  $p_{K_c}$  is a submersion, there are  $A, B \in M_c$  such that  $T_x p_{K_c} A = \bar{A}$ ,  $T_x p_{K_c} B = \bar{B}$ . We define  $\bar{\omega}_c(z)(\bar{A}, \bar{B}) = \omega_c(x)(A, B)$ . To prove that this is a good definition observe first that it is a consistent definition for fixed  $x$ , which is easy to prove, using the fact that  $\ker \omega_c(x) = \ker T_x p_{K_c}$ . Now choose a Darboux chart centered at  $x$ , say  $U \times V$ , such that, in this chart,  $p_{K_c}: U \times V \rightarrow U$  and  $\omega_c(x^1, x^2) = \bar{\omega}_c(x^1)$ , where  $\omega_c(x^1, x^2)$  and  $\bar{\omega}_c(x^1)$  are independent of  $(x_1, x_2)$ . This shows that  $\bar{\omega}_c$  is well defined on the chart. Using this and the fact that one can cover the connected submanifold  $p_{K_c}^{-1}(z)$  with charts as explained above, one can deduce by a simple argument that  $\bar{\omega}_c(z)$  is well defined.

(b) Let  $g$  be a Riemannian metric on  $M_c$ . Then for each  $x \in M_c$  there is a uniquely determined  $X(x) \in T_x M_c$  such that  $X(x)$  is orthogonal to  $\ker T_x p_{K_c}$  and  $T_x p_{K_c} X(x) = \bar{X}(x)$ , for all  $x \in M_c$ . This defines a vector field  $X$  on  $M_c$  which is  $p_{K_c}$ -related to  $\bar{X}$ .

(c) Given  $\bar{f}$  and using the result of (b) we see that there is a vector field  $X$  on  $M_c$  that is  $p_{K_c}$ -related to  $X_{\bar{f}}$ . Then, for every  $x \in M_c$  and every  $Y_x \in T_x M_c$ ,

$$\begin{aligned}\omega_c(x)(X(x), Y_x) &= \bar{\omega}_c(p_{K_c}(x))(X_{\bar{f}}(p_{K_c}(x)), T_x p_{K_c} Y_x) = d\bar{f}(p_{K_c}(x))(T_x p_{K_c} Y_x) \\ &= d(p_{K_c}^* \bar{f})(x)(Y_x).\end{aligned}$$

(d) One can proceed as in (b) and (c), choosing  $\bar{f}$  such that  $(d\bar{f}(p_{K_c}(x_0)))^\# = T_{x_0} p_{K_c} X_{x_0}$  and, besides, the metric  $g$  such that  $X_{x_0}$  is perpendicular to  $\ker T_{x_0} p_{K_c}$ .  $\blacksquare$

**Definition 3.10.** (a) For any subspace  $A \subseteq \mathcal{F}(N)$  define the distribution  $\Delta_A \subseteq TN$  by  $\Delta_A(x) = \{X_f(x) \mid f \in A\}$ .

(b) The space of **first class functions** is defined as

$$R^{(c)} = \{f \in \mathcal{F}(N) \mid X_f(x) \in T_x M_c, \text{ for all } x \in M_c\}.$$

In other words,  $R^{(c)}$  is the largest subset of  $\mathcal{F}(N)$  satisfying

$$\Delta_{R^{(c)}}(x) \subseteq T_x M_c,$$

$x \in M_c$ .

*Remark 3.11.* Dirac was interested in classical mechanics, where states are represented by points in phase space, as well as in quantum mechanics where this is not the case. From the point of view of classical mechanics, among the constraint submanifolds and constraints the only ones that seem to play an important role are  $M$ ,  $M_c$  and the constraints  $\phi_i^{(0)}$ ,  $\phi_i^{(c)}$  defining them by equations  $\phi_i^{(0)} = 0$ ,  $\phi_i^{(c)} = 0$ , respectively.

**Lemma 3.12.** (a)  $R^{(c)}$  is a Poisson subalgebra of  $(\mathcal{F}(N), \{, \})$ .

(b)  $M_c$  is an integral submanifold of  $\Delta_{R^{(c)}}$ . Moreover, for any vector field  $X$  on  $M_c$  that is  $p_{K_c}$ -related to a vector field  $X_{\bar{f}}$  on  $\bar{M}_c$  there exists a function  $f \in R^{(c)}$  such that  $f|_{M_c} = p_{K_c}^* \bar{f}$  and  $X = X_f|_{M_c}$ . In particular, any vector field  $X$  on  $M_c$  satisfying  $X(x) \in \ker \omega_c(x)$  for all  $x \in M_c$  is  $p_{K_c}$ -related to the vector field 0 on the symplectic manifold  $\bar{M}_c$ , which is associated to the function  $\bar{f} = 0$ , therefore there exists a function  $f \in R^{(c)}$ , which satisfies  $f|_{M_c} = 0$ , such that  $X(x) = X_f(x)$ ,  $x \in M_c$ .

**Proof.** (a) Let  $f, g \in R^{(c)}$ . Then  $X_f(x)$  and  $X_g(x)$  are both tangent to  $M_c$  at points  $x$  of  $M_c$  which implies that  $-X_{\{f, g\}}(x) = [X_f, X_g](x)$  is also tangent to  $M_c$  at points of  $x$  of  $M_c$ . This shows that  $\{f, g\} \in R^{(c)}$ . It is easy to see that any linear combination of  $f, g$  and also  $fg$  belong to  $R^{(c)}$ .

(b) By definition  $\Delta_{R^{(c)}} \subseteq TM_c$ . We need to show the converse inclusion. Let  $X_{x_0} \in T_{x_0} M_c$ , we need to find  $f \in R^{(c)}$  such that  $X_f(x_0) = X_{x_0}$ . Choose the function  $\bar{f}$  and the vector field  $X$  on  $M_c$  as in Lemma 3.9, (d). Choose any extension of  $p_{K_c}^* \bar{f}$  to a function  $g$  on  $N$ . For each  $x \in M_c$ , we can apply Lemmas 3.5 and 3.6 with  $E := T_x N$ ,  $V := T_x M_c$ ,  $\beta := dg(x)$ ,  $\gamma_i(x) := d\phi_i^{(c)}(x)$ ,  $i = 1, \dots, r_c$ . We obtain

that in a neighborhood  $U \subseteq N$  of each point  $x_0$  of  $M_c$  we can choose  $\mathcal{C}^\infty$  functions  $\lambda_{(c)}^i(x)$ ,  $i = 1, \dots, r_c$ , such that

$$X(x) = \left( dg(x) + \lambda_{(c)}^i(x) d\phi_i^{(c)}(x) \right)^\sharp,$$

for all  $x \in M_c \cap U$ . Let  $f_U(x) = g(x) + \lambda_{(c)}^i(x) \phi_i^{(c)}(x)$ , for all  $x \in U$ . Then we have that  $X_{f_U}(x) = X(x)$ , for all  $x \in M_c \cap U$ .

Now, consider a partition of unity  $\rho_i$ ,  $i \in I$ , on  $N$ , where each  $\rho_i$  is defined on an open set  $U_i$ ,  $i \in I$ . Let  $J \subseteq I$  be defined by the condition  $i \in J$  if and only if  $U_i \cap M_c \neq \emptyset$ . Using standard techniques of partitions of unity and the above result one can assume without loss of generality that for each  $i \in J$  there is a function  $f_{U_i}$  defined on  $U_i$  such that  $X_{f_{U_i}}(x) = X(x)$ , for all  $x \in M_c \cap U_i$ . Let  $f = \sum_{i \in J} \rho_i f_{U_i}$ , which can be naturally extended by 0 on  $N$ . Then it is easy to see, using the fact that  $f_{U_i}(x) = p_{K_c}^* \bar{f}(x) = g(x)$ , for each  $x \in M_c$ , that  $X(x) = X_f(x)$ , for each  $x \in M_c$ , and in particular  $X_0 = X(x_0) = X_f(x_0)$ . ■

**Lemma 3.13.** (a) *Each function  $f \in R^{(c)}$  is locally constant on the leaves of  $K_c$  therefore, since they are connected, for each  $f \in R^{(c)}$  there is a uniquely determined  $\bar{f} \in \mathcal{F}(\bar{M}_c)$ , called  $(p_{K_c})_* f$ , such that  $f|_{M_c} = p_{K_c}^* \bar{f}$ . Moreover, the vector fields  $X_f(x)$ ,  $x \in M_c$ , on  $M_c$  and  $X_{\bar{f}}$  on  $\bar{M}_c$  are  $p_{K_c}$ -related.*

(b) *For each  $\bar{f} \in \mathcal{F}(\bar{M}_c)$  there exists  $f \in R^{(c)}$  such that  $f|_{M_c} = p_{K_c}^* \bar{f}$  and the vector fields  $X_f(x)$ ,  $x \in M_c$ , on  $M_c$  and  $X_{\bar{f}}$  on  $\bar{M}_c$  are  $p_{K_c}$ -related.*

**Proof.** (a) Let  $f \in R^{(c)}$ , we only need to show that  $f$  is constant on the leaves of  $K_c$ , which is equivalent to showing that  $df(x)|\ker \omega_c(x) = 0$ , for all  $x \in M_c$ . For a given  $x \in M_c$ , let  $v_x \in \ker \omega_c(x)$ ; then using Lemma 3.12 (b) one sees that there is a function  $g \in R^{(c)}$  such that  $v_x = X_g(x)$ . Then we have

$$0 = \omega_c(x) (X_f(x), X_g(x)) = \Omega(x) (X_f(x), X_g(x)) = df(x) X_g(x).$$

Now we shall prove that  $X_f$  and  $X_{\bar{f}}$  are  $p_{K_c}$ -related. For each  $x \in M_c$  and each  $Y_x \in T_x M_c$ , we have

$$\omega_c(X_f(x), Y_x) = \Omega(X_f(x), Y_x) = df(x)(Y_x) = d(p_{K_c}^* \bar{f})(x)(Y_x).$$

Using this we obtain

$$\omega_c(x)(X_f(x), Y_x) = \bar{\omega}_c(x)(p_{K_c}(x))(T_x p_{K_c} X_f(x), T_x p_{K_c} Y_x) = d\bar{f}(p_{K_c}(x))(T_x p_{K_c} Y_x),$$

which shows that  $X_{\bar{f}}(p_{K_c}(x)) = T_x p_{K_c} X_f(x)$ , because  $\bar{\omega}_c$  is symplectic and  $T_x p_{K_c} Y_x$  represents an arbitrary element of  $T_{p_{K_c}(x)} \bar{M}$ .

(b) To find  $f$  we choose a vector field  $X$  that is  $p_{K_c}$ -related to  $X_{\bar{f}}$  according to Lemma 3.9 and then use Lemma 3.12 (b). ■

**Definition 3.14.**

$$I^{(c)} = \{f \in R^{(c)} \mid f|_{M_c} = 0\},$$

$$Z_{I^{(c)}} R^{(c)} = \{f \in R^{(c)} \mid \{f, h\} \in I^{(c)}, \text{ for all } h \in R^{(c)}\}.$$

Elements of  $I^{(c)}$  are called **first class constraints**.



**Lemma 3.15.** (a)  $I^{(c)}$  is a Poisson ideal of  $R^{(c)}$ , that is, it is an ideal of the ring  $R^{(c)}$  such that if  $f \in I^{(c)}$ , then  $\{f, h\} \in I^{(c)}$ , for all  $h \in R^{(c)}$ .

(b)  $Z_{I^{(c)}} R^{(c)}$  is a Poisson subalgebra of  $R^{(c)}$ .

**Proof.** (a) Let  $f, g \in I^{(c)}$  and  $h \in R^{(c)}$ . Then it is immediate that  $f + g$  and  $hg$  belong to  $I^{(c)}$ . For any  $h \in R^{(c)}$ , we have  $\{f, h\}|_{M_c} = X_h(f)|_{M_c} = 0$ .

(b) Follows from (a), using basic Poisson algebra arguments. ■

**Lemma 3.16.** The following conditions are equivalent for a function  $f \in R^{(c)}$ .

(i)  $f \in Z_{I^{(c)}} R^{(c)}$ .

(ii)  $f|_{M_c}$  is locally constant.

(iii)  $X_f(x) \in \ker \omega_c(x)$  for  $x \in M_c$ .

**Proof.** Assume (i). Then  $\{f, h\}|_{M_c} = 0$  for all  $h \in R^{(c)}$ , that is,  $df(x)X_h(x)|_{M_c} = 0$ . By Lemma 3.12, (b), we know that  $X_h(x)$  represents any vector in  $T_x M_c$ . We can conclude that  $f|_{M_c}$  is locally constant, so (ii) holds. Now we will prove that (ii) implies (iii). Let  $f|_{M_c}$  be locally constant. Then for all  $g \in R^{(c)}$  and all  $x \in M_c$ ,

$$0 = X_g(f)(x) = \Omega(x)(X_f, X_g)(x) = \omega_c(x)(X_f(x), X_g(x)).$$

Since, again by Lemma 3.12,  $X_g(x)$  represents any element of  $T_x M_c$ , we can conclude that  $X_f(x) \in \ker \omega_c(x)$ , so (iii) holds true. Now we will prove that (iii) implies (i). Assume that  $X_f(x) \in \ker \omega_c(x)$ ,  $x \in M_c$ . Then for all  $g \in R^{(c)}$  and all  $x \in M_c$ ,

$$\{g, f\}(x) = \Omega(x)(X_g, X_f)(x) = \omega_c(x)(X_g(x), X_f(x)) = 0,$$

that is,  $\{g, f\} \in I^{(c)}$ . Using this and the definitions, we see that  $f \in Z_{I^{(c)}} R^{(c)}$ . ■

**Lemma 3.17.** The map  $(p_{K_c})_*: R^{(c)} \rightarrow \mathcal{F}(\bar{M}_c)$  defined in Lemma 3.13 is a surjective Poisson map and its kernel is  $I^{(c)}$ , therefore there is a natural isomorphism of Poisson algebras  $(p_{K_c})_{*I^{(c)}}: R^{(c)}/I^{(c)} \rightarrow \mathcal{F}(\bar{M}_c)$ .

**Proof.** Surjectivity of  $(p_{K_c})_*$  and the fact that its kernel is  $I^{(c)}$  follows immediately from Lemma 3.13 and the definitions. This implies that  $(p_{K_c})_{*I^{(c)}}$  is an algebra isomorphism. Also, using the definitions, for  $f, g \in R^{(c)}$  and any  $x \in M_c$  we can prove easily that

$$\begin{aligned} \{f, g\}(x) &= \Omega(x)(X_f(x), X_g(x)) = \omega_c(x)(X_f(x), X_g(x)) \\ &= \bar{\omega}_c(p_{K_c}(x))(X_{\bar{f}}(p_{K_c}(x)), X_{\bar{g}}(p_{K_c}(x))) = \{\bar{f}, \bar{g}\}(p_{K_c}(x)), \end{aligned}$$

where  $\bar{f} = (p_{K_c})_* f$ ,  $\bar{g} = (p_{K_c})_* g$ . Denote by  $\pi_{I^{(c)}}: R^{(c)} \rightarrow R^{(c)}/I^{(c)}$  the natural homomorphism of Poisson algebras. Then from the previous equalities we obtain  $(p_{K_c})_{*I^{(c)}}\{\pi_{I^{(c)}}(f), \pi_{I^{(c)}}(g)\} = \{\bar{f}, \bar{g}\}$ , which shows that  $(p_{K_c})_{*I^{(c)}}$  is a Poisson isomorphism. In other words, we have the commutative diagram

$$\begin{array}{ccc} R^{(c)} & \xrightarrow{\pi_{I^{(c)}}} & R^{(c)}/I^{(c)} \\ & \searrow (p_{K_c})_* & \downarrow (p_{K_c})_{*I^{(c)}} \\ & & \mathcal{F}(\bar{M}_c) \end{array}$$

All the arrows are defined in a natural way and they are surjective Poisson algebra homomorphisms. ■

**Physical variables.** It is immediate to see from the definitions that for all  $x \in M_c$ ,

$$\ker \omega(x) \cap T_x M_c \subseteq \ker \omega_c(x). \quad (3.5)$$

From now on we will assume the following.

**Assumption 3.18.** (a)  $\ker \omega(x)$  is a regular distribution, that is, it determines a regular foliation  $K$  and the natural projection  $p_K: M \rightarrow \bar{M}$ , where  $\bar{M} = M/K$ , is a submersion.

(b)  $\ker \omega(x) \cap T_x M_c$  is a distribution of constant rank.

**Theorem 3.19.** The distribution  $\ker \omega(x) \cap T_x M_c$  is regular and has rank  $d^{(c)}(x)$ . Its integral manifolds are  $S \cap M_c$ , where  $S$  is an integral manifold of  $\ker \omega$ . Moreover, these integral manifolds give a foliation  $\tilde{K}_c$  of  $M_c$  which is regular, that is, the natural map  $p_{\tilde{K}_c}: M_c \rightarrow \tilde{M}_c$ , where  $\tilde{M}_c = M_c/\tilde{K}_c$  is a submersion. Besides, each leaf of the foliation  $K_c$  is foliated by leaves of  $\tilde{K}_c$ , which gives a naturally defined submersion  $p_{K_c \tilde{K}_c}: \tilde{M}_c \rightarrow \bar{M}_c$ . In other words, we obtain the commutative diagram

$$\begin{array}{ccc} M_c & \xrightarrow{p_{\tilde{K}_c}} & \tilde{M}_c \\ & \searrow p_{K_c} & \downarrow p_{K_c \tilde{K}_c} \\ & & \bar{M}_c \end{array}$$

where each arrow is a naturally defined submersion.

**Proof.** The first assertion, about the rank of the distribution  $\ker \omega(x) \cap T_x M_c$ , is easy to prove. Let  $x_0 \in M_c$ . Then there exists a uniquely determined integral manifold  $S$  of the distribution  $\ker \omega$  such that  $x_0 \in S$ . Using that, by assumption,  $\ker \omega(x) \cap T_x M_c$  is a distribution of constant dimension and that  $\dim(\ker \omega(x) \cap T_x M_c) = \dim(T_x S \cap T_x M_c)$  we can conclude that the intersection  $S \cap M_c$  coincides with the integral leaf of the integrable distribution of  $\ker \omega \cap TM_c$  containing  $x_0$ . So we obtain the foliation  $\tilde{K}_c$  of  $M_c$ . Using (3.5) we can deduce that each leaf of the foliation  $K_c$  is foliated by leaves of  $\tilde{K}_c$ . The rest of the proof follows by standard arguments. ■

**Lemma 3.20.** (a) The following diagram is commutative

$$\begin{array}{ccccccc} M & \xrightarrow{p_K} & \bar{M} & \xleftarrow{\tau_{\bar{M}}} & T\bar{M} & & \\ f_c \uparrow & & \uparrow \tilde{f}_c & & \uparrow \tilde{F}_c \circ \epsilon_c & \nwarrow \tilde{F}_c & \\ M_c & \xrightarrow{p_{\tilde{K}_c}} & \tilde{M}_c & \xleftarrow{\tau_{\tilde{M}}} & T\bar{M}|_{\tilde{M}_c} & \xrightarrow{\epsilon_c} & \tilde{f}_c^* T\bar{M} \\ & \searrow p_{K_c} & \downarrow p_{K_c \tilde{K}_c} & & & & \\ & & \bar{M}_c & & & & \end{array}$$

where the arrows are defined as follows. The maps  $p_K$ ,  $p_{K_c}$ ,  $p_{\tilde{K}_c}$  and  $p_{K_c \tilde{K}_c}$  are defined in Assumption 3.8, Assumption 3.18 and Theorem 3.19. By definition, the map  $f_c$  is the inclusion. The map  $\tilde{f}_c$  is an embedding defined by  $\tilde{f}_c(S \cap M_c) = S$ ,

where  $S$  is a leaf of the foliation  $K$ . We will think of  $\tilde{f}_c$  as being an inclusion. The vector bundle  $T\bar{M}|_{\bar{M}_c}$  is the tangent bundle  $T\bar{M}$  restricted to  $\bar{M}_c$ . In other words, since  $\tilde{f}_c$  is an inclusion,  $T\bar{M}|_{\bar{M}_c}$  is identified via some isomorphism, called  $\epsilon_c$ , with the pullback of  $T\bar{M}$  by  $\tilde{f}_c$ . We call  $\tilde{F}_c$  the natural map associated to the pullback.

(b) The presymplectic form  $\omega$  on  $M$  passes to the quotient via  $p_K$  giving a uniquely defined symplectic form  $\bar{\omega}$  on  $\bar{M}$ , satisfying  $p_K^* \bar{\omega} = \omega$ . The presymplectic form  $\omega_c$ , which, by definition is  $f_c^* \omega$ , defines uniquely a presymplectic form  $\tilde{\omega}_c$  on  $\bar{M}_c$  via  $p_{\tilde{K}_c}^* \tilde{\omega}_c = \omega_c$ ,  $\tilde{\omega}_c = \tilde{f}_c^* \bar{\omega}$ . The energy  $\mathcal{E}$  on  $M$  satisfies  $d\mathcal{E}(x)|_{\ker \omega(x)} = 0$ , for all  $x \in M_c$ , therefore it defines uniquely a 1-form on  $T\bar{M}|_{\bar{M}_c}$ , called  $(\tilde{F}_c \circ \epsilon)^* d\mathcal{E} \in \Gamma((T\bar{M}|_{\bar{M}_c})^*)$ . Since  $\mathcal{E}$  is constant on each leaf of  $\tilde{K}_c$ , it also defines a function  $\tilde{\mathcal{E}}_c$  on  $\bar{M}_c$ . Since  $T\bar{M}_c \subseteq T\bar{M}$  via the inclusion  $T\tilde{f}_c$  we have  $(\tilde{F}_c \circ \epsilon)^* d\mathcal{E}|_{T\bar{M}_c} = d\tilde{\mathcal{E}}_c(\tilde{x})$ , for all  $\tilde{x} \in \bar{M}_c$ .

(c) Equation of motion (3.4) on  $M_c$  passes to the quotient  $\bar{M}_c$  as

$$\bar{\omega}(\tilde{x})(\tilde{X}(\tilde{x}), \cdot) = (\tilde{F}_c \circ \epsilon)^* d\mathcal{E}(\tilde{x}), \quad (3.6)$$

where  $\tilde{X}(\tilde{x}) \in T_{\tilde{x}}\bar{M}_c$ . This means that if  $X(x) \in T_x M_c$  is a solution of (3.4) then  $\tilde{X}(\tilde{x}) := T_x p_{\tilde{K}_c} X(x)$ , where  $\tilde{x} = p_{\tilde{K}_c}(x)$ , is a solution of (3.6). Therefore, a solution curve  $x(t)$  of (3.4) projects to a solution curve  $\tilde{x}(t) = p_{\tilde{K}_c}(x(t))$  of (3.6) on  $\bar{M}_c$ . Equation (3.6) has unique solution  $\tilde{X}(\tilde{x})$  for each  $\tilde{x} \in \bar{M}_c$ . This solution also satisfies the equation

$$\tilde{\omega}_c(\tilde{x})(\tilde{X}(\tilde{x}), \cdot) = d\tilde{\mathcal{E}}(\tilde{x}). \quad (3.7)$$

However solutions to equation (3.7) are not necessarily unique, since  $\ker \tilde{\omega}_c(\tilde{x})$  is not necessarily 0.

(d) The restriction of the energy  $\mathcal{E}|_{M_c}$  satisfies

$$d(\mathcal{E}|_{M_c})(x)|_{\ker \omega_c(x)} = 0,$$

for all  $x \in M_c$ , therefore there is a uniquely defined function  $\bar{\mathcal{E}}_c$  on  $\bar{M}_c$  such that  $p_{\tilde{K}_c}^* \bar{\mathcal{E}}_c = \mathcal{E}|_{M_c}$ . The equation

$$\bar{\omega}_c(\bar{x})(\bar{X}(\bar{x}), \cdot) = d\bar{\mathcal{E}}_c(\bar{x}) \quad (3.8)$$

has unique solution  $\bar{X}(\bar{x})$  for  $\bar{x} \in \bar{M}_c$ . If  $\tilde{X}(\tilde{x})$  is a solution of (3.6) then  $\bar{X}(\bar{x}) = T_{\tilde{x}} p_{\tilde{K}_c \tilde{K}_c} \tilde{X}(\tilde{x})$  is a solution of (3.8). Therefore, a solution curve  $\tilde{x}(t)$  of (3.6) projects to a solution curve  $\bar{x}(t) = p_{\tilde{K}_c \tilde{K}_c}(\tilde{x}(t))$  of (3.8) on  $\bar{M}_c$ .

**Proof.** (a) The equality  $p_{K_c} = p_{K_c \tilde{K}_c} \circ p_{\tilde{K}_c}$  was proven in Theorem 3.19. The equality  $\tau_{\bar{M}} \circ \tilde{F}_c \circ \epsilon_c = \tilde{f}_c \circ \tau_{\bar{M}}$  results immediately from the definitions. The equality  $p_K \circ f_c = \tilde{f}_c \circ p_{\tilde{K}_c}$  results by applying the definitions and showing that, for given  $x \in M_c$ ,  $p_K \circ f_c(x) = S_x = \tilde{f}_c \circ p_{\tilde{K}_c}(x)$ , where  $S_x$  is the only leaf of  $K$  containing  $x$ .

(b) Existence and uniqueness of  $\bar{\omega}$  and  $\tilde{\omega}$  is a direct consequence of the definitions and standard arguments on passing to quotients. For any  $x \in M_c$  we know that there exists a solution  $X$  of equation (3.4), from which it follows immediately that

$d\mathcal{E}(x)|\ker\omega(x) = 0$ . The rest of the proof consists of standard arguments on passing to quotients.

(c) We shall omit the proof of this item which is a direct consequence of the definitions and standard arguments on passing to quotients.

(d) If  $X(x)$  is a solution of  $\omega(x)(X(x), \cdot) = d\mathcal{E}(x)$  then it is clear that it also satisfies  $\omega_c(x)(X(x), \cdot) = d(\mathcal{E}|M_c)(x)$ . It follows that

$$d(\mathcal{E}|M_c)(x)|\ker\omega_c(x) = 0,$$

for all  $x \in M_c$ . The rest of the proof is a consequence of standard arguments on passing to quotients.  $\blacksquare$

*Remark 3.21.* Recall that the locally constant function  $d^{(c)}(x)$  on  $M_c$  is the dimension of the distribution  $\ker\omega(x) \cap T_x M_c$  on  $M_c$  and also the dimension of the fiber of the bundle  $S^{(c)}$ . If  $d^{(c)}(x)$  is nonzero then there is no uniqueness of solution to equation of motion (3.4), since solution curves to that equation satisfy, by definition,

$$\begin{aligned} \omega(x)(\dot{x}, \cdot) &= d\mathcal{E}(x)|T_x M \\ (x, \dot{x}) &\in T_x M_c. \end{aligned}$$

Passing to the quotient manifold  $\widetilde{M}_c$  eliminates this indeterminacy and uniqueness of solution is recovered. This is related to the notions of *physical state* and *determinism*, mentioned by Dirac [1964, p. 20].

The reduced equations of motion on the symplectic manifold  $\widetilde{M}_c$  are Hamilton's equations. The reduced equations of motion on the manifold  $\widetilde{M}_c$  are not necessarily Hamilton's equations but they are ODEs. The natural fibration  $p_{K_c \widetilde{K}_c} : \widetilde{M}_c \rightarrow \bar{M}_c$  has the following property. For  $\bar{m}_0 \in \bar{M}_c$  and  $\tilde{m}_0 \in (p_{K_c \widetilde{K}_c})^{-1}(\bar{m}_0) \subseteq \widetilde{M}_c$ , if  $\bar{m}(t)$  and  $\tilde{m}(t)$  are the unique solution curves through  $\bar{m}_0$  and  $\tilde{m}_0$  respectively, then  $p_{K_c \widetilde{K}_c} \tilde{m}(t) = \bar{m}(t)$ . This means that there is a one-to-one correspondence between the solution curves on  $\widetilde{M}_c$  and  $\bar{M}_c$ , modulo the choice of a point on the fiber over the physical state  $\bar{m}_0$ . See also remark 3.30.

### 3.3 Second class subbundles

The notions of first class and second class constraints and functions and also second class submanifolds only depend on the final constraint submanifold  $M_c$  and the ambient symplectic manifold  $N$  and do not depend on the primary constraint  $M_0 = M$  or the energy  $\mathcal{E} : N \rightarrow \mathbb{R}$ . Accordingly, in this section we will adopt an abstract setting, where only an ambient symplectic manifold and a submanifold are given. Among several interesting references we cite Śniatycki [1974] which has some points of contact with our work.

The following definition is inspired by the ones given in section 3.2.

**Definition 3.22.** Let  $(P, \Omega)$  be a symplectic manifold and  $S \subseteq P$  a given submanifold. Then, by definition,

$$\begin{aligned} R^{(S,P)} &:= \{f \in \mathcal{F}(P) \mid X_f(x) \in T_x S, \text{ for all } x \in S\} \\ I^{(S,P)} &:= \{f \in R^{(S,P)} \mid f|_S = 0\} \end{aligned}$$

Elements of  $R^{(S,P)}$  are called **first class functions**. Elements of  $I^{(S,P)}$  are called **first class constraints**. The submanifold  $S$  is called a **first class constraint submanifold** if for all  $f \in \mathcal{F}(P)$  the condition  $f|_S = 0$  implies  $f \in I^{(S,P)}$ , that is,  $I_{(S,P)} \subseteq I^{(S,P)}$ , where  $I_{(S,P)}$  is the ideal of the ring  $\mathcal{F}(P)$  of all functions vanishing on  $S$ .

Obviously, using the notation introduced before in the present section,  $R^{(M_c,N)} = R^{(c)}$  and  $I^{(M_c,N)} = I^{(c)}$ . All the properties proven for  $R^{(c)}$  and  $I^{(c)}$  hold in general for  $R^{(S,P)}$  and  $I^{(S,P)}$ . For instance,  $T_x S$  is the set of all  $X_f(x)$ ,  $f \in R^{(S,P)}$ . Every function  $f \in R^{(S,P)}$  satisfies  $df(x)(X_x) = 0$ , for all  $X_x \in \ker(\Omega(x)|_{T_x S})$  and  $\ker(\Omega(x)|_{T_x S})$  is the set of all  $X_f(x)$ ,  $f \in I^{(S,P)}$ . We shall collect in a single lemma 3.23 the basic properties of first and second class constraints and Dirac brackets and their proofs in a compact way and with a uniform language that does not rely on coordinates. Most of these results are essentially known since the fundamental works of Dirac, and have been stated and proven in different ways in the literature, except perhaps for the notion of second class subbundle and the corresponding description of second class submanifolds as tangent to the subbundle. The second class subbundles  $V^\Omega \subseteq TP|_S$  classify all second class submanifolds  $S^V$  containing  $S$  at a linear level, that is,  $V^\Omega$  is tangent to the second class submanifold. For such a second class submanifold, which is a symplectic submanifold, the Dirac bracket of two functions  $F$  and  $G$  at points  $x \in S$  can be calculated as the canonical bracket of the restrictions of  $F$  and  $G$ . This has a global character. A careful study of the global existence of a bracket defined on sufficiently small open sets  $U \subseteq P$  containing  $S$  which coincides with the previous one on the second class submanifold will not be considered in this paper. However, to write global equations of motion on the final constraint submanifold one only needs to know the vector bundle  $V^\Omega$ , which carries a natural fiberwise symplectic form.

All these are fundamental properties of second class constraints and constraint submanifolds, and Theorem 3.24 collects the essential part of them; we suggest to take a look at it before reading Lemma 3.23.

**Lemma 3.23.** *Let  $(P, \Omega)$  be a symplectic manifold of dimension  $2n$  and  $S \subseteq P$  a given submanifold of codimension  $r$ . Let  $\omega$  be the pullback of  $\Omega$  to  $S$  and assume that  $\ker \omega(x)$  has constant dimension. Assume that  $S$  is defined regularly by equations  $\phi_1 = 0, \dots, \phi_r = 0$  on a neighborhood  $U \supseteq S$  and assume that we can choose a subset  $\{\chi_1, \dots, \chi_{2s}\} \subseteq \{\phi_1, \dots, \phi_r\}$  such that  $\det(\{\chi_i, \chi_j\}(x)) \neq 0$  for all  $x \in U$ , where we assume that  $2s = \text{rank}(\{\phi_i, \phi_j\}(x))$ , for all  $x \in U$ . We shall often denote  $c_{ij}^X(x) = \{\chi_i, \chi_j\}(x)$  and  $c_{ij}^{X^{-1}}(x)$  the inverse of  $c_{ij}^X(x)$ . Moreover, we will assume that the following stronger condition holds, for simplicity. Equations  $\phi_1 = B_1, \dots, \phi_r = B_r$  and  $\chi_1 = C_1, \dots, \chi_{2s} = C_{2s}$  define submanifolds of  $U$  regularly, for small enough  $B_1, \dots, B_r$  and  $C_1, \dots, C_{2s}$ . Then*

(a)  $2s = r - \dim \ker \omega = 2n - \dim S - \dim \ker \omega$ . There are  $\psi_k \in I^{(S,P)}$ ,  $k = 1, \dots, r - 2s$ , which in particular implies  $\{\psi_k, \psi_l\}(x) = 0$ ,  $\{\psi_k, \chi_i\}(x) = 0$ , for  $k, l = 1, \dots, r - 2s$ ,  $i = 1, \dots, 2s$ , such that  $d\psi_1(x), \dots, d\psi_{r-2s}(x), d\chi_1(x), \dots, d\chi_{2s}(x)$ , are linearly independent for all  $x \in S$ . Moreover,  $X_{\psi_1}(x), \dots, X_{\psi_{r-2s}}(x)$  form a basis of  $\ker \omega(x)$ , for all  $x \in S$  and  $d\psi_1(x), \dots, d\psi_{r-2s}(x), d\chi_1(x), \dots, d\chi_{2s}(x)$  form a basis

of  $(T_x S)^\circ$ .

(b) The vector subbundle  $V^\chi \subseteq TP|S$  with base  $S$  and fiber

$$V_x^\chi = \text{span}(X_{\chi_1}(x), \dots, X_{\chi_{2s}}(x)) \subseteq T_x P,$$

satisfies

$$V_x^\chi \cap T_x S = \{0\} \quad (3.9)$$

$$V_x^\chi \oplus \ker \omega(x) = (T_x S)^\Omega \quad (3.10)$$

$$(V_x^\chi)^\Omega \cap (\ker \omega(x))^\Omega = T_x S, \quad (3.11)$$

$x \in S$ .

(c) There is a neighborhood  $U$  of  $S$  such that the equations  $\chi_1 = 0, \dots, \chi_{2s} = 0$ , on  $U$  define a symplectic submanifold  $S^\chi$  such that  $S \subseteq S^\chi$  and

$$\begin{aligned} T_x S^\chi &= (V_x^\chi)^\Omega \\ T_x S^\chi \oplus V_x^\chi &= T_x P, \end{aligned}$$

for  $x \in S^\chi$ , where we have extended the definition of  $V_x^\chi$  for  $x \in S^\chi$  using the expression

$$V_x^\chi = \text{span}(X_{\chi_1}(x), \dots, X_{\chi_{2s}}(x)) \subseteq T_x P,$$

for  $x \in S^\chi$ . The submanifold  $S^\chi$  has the property  $I_{(S, S^\chi)} \subseteq I^{(S, S^\chi)}$ , that is,  $S$  is a first class constraint submanifold of  $S^\chi$ , defined regularly by  $\psi_i|_{S^\chi} = 0$ ,  $i = 1, \dots, r - 2s$ , and  $\psi_i|_{S^\chi} \in I^{(S, S^\chi)}$ ,  $i = 1, \dots, r - 2s$ . Moreover, it has the only possible dimension, which is  $\dim S^\chi = \dim S + \dim \ker \omega = 2n - 2s$ , for symplectic submanifolds having that property. It is also a minimal object in the set of all symplectic submanifolds  $P_1 \subseteq P$ , ordered by inclusion, satisfying  $S \subseteq P_1$ .

(d) Let  $V$  be any vector subbundle of  $TP|S$  such that

$$V \oplus \ker \omega = (TS)^\Omega, \quad (3.12)$$

or equivalently,

$$V^\flat \oplus (\ker \omega)^\flat = (TS)^\circ. \quad (3.13)$$

Then  $\dim V_x = 2s$ , for  $x \in S$ . Let  $S^V$  be a submanifold of  $P$  such that  $T_x S^V = V_x^\Omega$ , for each  $x \in S$ . Then  $S$  is a submanifold of  $S^V$ . Such a submanifold  $S^V$  always exists. Moreover, for such a submanifold there is an open set  $U \subseteq P$  containing  $S$  such that  $S^V \cap U$  is a symplectic submanifold of  $P$ .

Let  $\bar{x} \in S$  and let  $\chi'_1 = 0, \dots, \chi'_{2s} = 0$  be equations defining  $S^V \cap U'$  for some open neighborhood  $U' \subseteq P$  and satisfying that  $d\chi'_1(x), \dots, d\chi'_{2s}(x)$  are linearly independent for  $x \in S^V \cap U'$ . Then,  $d\chi'_1(x), \dots, d\chi'_{2s}(x), d\psi_1(x), \dots, d\psi_{r-2s}(x)$  are linearly independent and  $\det(\{\chi'_i, \chi'_j\}(x)) \neq 0$ , for  $x \in S^V \cap U'$ . All the properties established in (a), (b), (c) for  $\chi_1, \dots, \chi_{2s}$  on  $S$  hold in an entirely similar way for  $\chi'_1, \dots, \chi'_{2s}$ , on  $S \cap U'$ . In particular,  $S^V \cap U' = S^{\chi'}$ .

(e) Let  $\omega^\chi$  be the pullback of  $\Omega$  to  $S^\chi$  and  $\{, \}_\chi$  the corresponding bracket. For given  $F, G \in \mathcal{F}(P)$  define  $F_\chi := F - \chi_i c_\chi^{ij} \{\chi_j, F\}$  and also

$$\{F, G\}_{(\chi)} := \{F, G\} - \{F, \chi_i\} c_\chi^{ij} \{\chi_j, G\},$$

which is the famous bracket introduced by Dirac, called **Dirac bracket**, and it is defined for  $x$  in the neighborhood  $U$  where  $c_{ij}^x(x)$  has an inverse  $c_{\chi}^{ij}(x)$ . Then, for any  $x \in S^x$ ,

$$\{F_{\chi}, \chi_i\}(x) = 0$$

for  $i = 1, \dots, 2s$ , and also

$$\{F_{\chi}, G_{\chi}\}(x) = \{F, G\}_{(\chi)}(x) = \omega^x(x) (X_{F|S^x}(x), X_{G|S^x}(x)) = \{F|S^x, G|S^x\}_{\chi}(x).$$

If we denote  $X_{(\chi), F}$  the Hamiltonian vector field associated to the function  $F \in \mathcal{F}(P)$ , with respect to the Dirac bracket  $\{, \}_{(\chi)}$  then the previous equalities are equivalent to

$$X_{(\chi), F}(x) = X_{F_{\chi}}(x) = X_{F|S^x}(x).$$

The Jacobi identity is satisfied for the Dirac bracket  $\{F, G\}_{(\chi)}$  on  $S^x$ , that is,

$$\{\{F, G\}_{(\chi)}, H\}_{(\chi)}(x) + \{\{H, F\}_{(\chi)}, G\}_{(\chi)}(x) + \{\{G, H\}_{(\chi)}, F\}_{(\chi)}(x) = 0,$$

for  $x \in S^x$ .

(f) Let  $U$  be an open neighborhood of  $S$  such that  $c_{ij}^x(x)$  is invertible for  $x \in U$ . For each  $C = (C_1, \dots, C_{2s}) \in \mathbb{R}^{2s}$  let  $\chi_i^C = \chi_i - C_i$  and define  $S^{x^C}$  by the equations  $\chi_i^C(x) = 0$ ,  $i = 1, \dots, 2s$ ,  $x \in U$ . For any  $C$  in a sufficiently small neighborhood of 0,  $S^{x^C}$  is a nonempty symplectic submanifold of  $P$ . Define the matrix  $c_{ij}^{x^C}(x) = \{\chi_i^C, \chi_j^C\}(x)$ , and also  $c_{\chi^C}^{ij}(x)$  as being its inverse,  $x \in U$ . Then, the equalities

$$c_{ij}^{x^C}(x) = \{\chi_i, \chi_j\}(x) = \{\chi_i^C, \chi_j^C\}(x) = c_{ij}^{x^C}(x), \quad (3.14)$$

and also,

$$\{F, G\}_{(\chi)}(x) = \{F, G\}_{(\chi^C)}(x) \quad (3.15)$$

are satisfied for all  $x \in U$ . All the definitions and properties proved in (e) for the case  $C = 0$  hold in general for any  $C$  in a neighborhood of 0 small enough to ensure that  $S^{x^C}$  is nonempty. In particular, the equalities

$$\begin{aligned} \{F_{\chi^C}, G_{\chi^C}\}(x) &= \{F, G\}_{(\chi^C)}(x) = \omega^{x^C}(x) (X_{F|S^{x^C}}(x), X_{G|S^{x^C}}(x)) \\ &= \{F|S^{x^C}, G|S^{x^C}\}_{\chi^C}(x) \end{aligned} \quad (3.16)$$

and

$$X_{(\chi^C), F}(x) = X_{F_{\chi^C}}(x) = X_{F|S^{x^C}}(x) \quad (3.17)$$

hold for  $x \in S^{x^C}$ , and any  $C$  in such a neighborhood. The Dirac bracket  $\{F, G\}_{(\chi)}$  satisfies the Jacobi identity for  $F, G \in \mathcal{F}(U)$  and the symplectic submanifolds  $S^{x^C}$  are the symplectic leaves of the Poisson manifold  $(\mathcal{F}(U), \{, \}_{(\chi)})$ . By shrinking, if necessary, the open set  $U$  and for  $C$  in a sufficiently small neighborhood of 0 in  $\mathbb{R}^{2s}$ , the equations  $\psi_k|S^{x^C} = 0$ ,  $k = 1, \dots, r - 2s$ , define regularly a first class constraint submanifold  $S^C \subseteq S^{x^C} \subseteq U$ , and the functions  $\psi_k|S^{x^C} \in R^{(S^C, S^{x^C})} \subseteq \mathcal{F}(S^{x^C})$  are first class constraints, that is,  $\psi_k|S^{x^C} \in I^{(S^C, S^{x^C})}$   $k = 1, \dots, r - 2s$ . We have that  $\dim S^{x^C} = \dim S^C + \dim \ker \omega^C$ , where  $\omega^C$  is the pullback of  $\Omega$  to  $S^C$ . One has  $\dim S^C = \dim S$  and  $\dim S^{x^C} = \dim S^x$ , therefore  $\dim \ker \omega = \dim \ker \omega^C$ .



**Proof. (a)** Let  $x \in S$ . We are going to use lemmas and corollaries 3.1–3.7 with  $E := T_x P$ ;  $V := T_x S$ ;  $\gamma_i := d\phi_i(x)$ ,  $i = 1, \dots, r$ ;  $\Omega := \Omega(x)$ ;  $\omega := \omega(x)$ ;  $\beta = 0$ .

Elements  $\psi = \lambda^i \phi_i$ ,  $\lambda^i \in \mathcal{F}(P)$  such that  $\psi \in I^{(S,P)}$ , which implies  $X_\psi(x) \in T_x S$  for  $x \in S$ , must satisfy  $\{\psi, \phi_j\}(x) = 0$ , or, equivalently,  $\lambda^i(x) d\phi_i(x) (X_{\phi_j}(x)) = 0$ , for  $j = 1, \dots, r$ ,  $x \in S$ . Using Lemma 3.6 we see that  $X_\psi(x) \in \ker \omega(x)$ . Since one can choose  $r - 2s$  linearly independent solutions, say  $\lambda_i = (\lambda_i^1, \dots, \lambda_i^r)$ ,  $i = 1, \dots, r - 2s$ , we obtain elements  $\psi_i \in I^{(S,P)}$ , namely,  $\psi_i = \lambda_i^j \phi_j$ , such that  $(d\psi_1(x), \dots, d\psi_{r-2s}(x))$  are linearly independent, or, equivalently, taking into account Corollary 3.7, that  $(X_{\psi_1}(x), \dots, X_{\psi_{r-2s}}(x))$  is a basis of  $\ker \omega(x)$  for  $x \in S$ . If  $d\psi_1(x), \dots, d\psi_{r-2s}(x), d\chi_1(x), \dots, d\chi_{2s}(x)$ , were not linearly independent, then there would be a linear combination, say  $\bar{\chi} = a^i \chi_i$ , with at least one nonzero coefficient, and some  $x \in S$ , such that  $d\bar{\chi}(x) = \mu^k d\psi_k(x)$  for some  $\mu^k$ ,  $k = 1, \dots, r - 2s$ . But then, for any  $j = 1, \dots, 2s$ ,  $\{\bar{\chi}, \chi_j\}(x) = d\bar{\chi}(x) X_{\chi_j}(x) = \mu^k d\psi_k(x) X_{\chi_j}(x) = -\mu^k d\chi_j(x) X_{\psi_k}(x) = 0$ , which contradicts the fact that  $\det(\{\chi_i, \chi_j\}(x)) \neq 0$ . Using this and the fact that  $\psi_i = 0$ ,  $\chi_j = 0$ ,  $i = r - 2s, j = 1, \dots, 2s$  define  $S$  regularly, we can conclude that  $d\psi_1(x), \dots, d\psi_{r-2s}(x), d\chi_1(x), \dots, d\chi_{2s}(x)$  is a basis of  $(T_x S)^\circ$ .

**(b)** If  $\lambda^i X_{\chi_i}(x) \in T_x S$  then  $d\chi_j(x) \lambda^i X_{\chi_i}(x) = 0$ ,  $j = 1, \dots, 2s$ , which implies  $\lambda^i \{\chi_j, \chi_i\} = 0$ ,  $j = 1, \dots, 2s$ , then  $\lambda^i = 0$ ,  $i = 1, \dots, 2s$ , which proves (3.9). To prove (3.10) we apply the operator  $^b$  to both sides and obtain the equivalent equality  $\text{span}(d\chi_1(x), \dots, d\chi_{2s}(x)) \oplus \text{span}(d\psi_1(x), \dots, d\psi_{r-2s}(x)) = (T_x S)^\circ$ , which we know is true, as proven in (a). To prove (3.11) we apply the orthogonal operator  $^\Omega$  to both sides of (3.10).

**(c)** Since  $d\chi_1(x), \dots, d\chi_{2s}(x)$  are linearly independent for  $x \in S$  they are also linearly independent for  $x$  in a certain neighborhood  $U$  of  $S$ . Then  $\chi_1(x) = 0, \dots, \chi_{2s}(x) = 0$  define a submanifold  $S^\chi$  of  $U$  containing  $S$ . To see that it is a symplectic submanifold choose  $x \in S^\chi$  and apply Corollary 3.7 with  $E := T_x P$ ;  $V := T_x S^\chi$ ;  $\beta := 0$ ;  $\gamma_i := d\chi_i$ ,  $i = 1, \dots, 2s$ ;  $\omega := \Omega(x)|_{T_x S^\chi}$ . We can conclude that  $\dim(\ker \Omega(x)|_{T_x S^\chi}) = 0$ . Now let us prove that  $T_x S^\chi = (V_x^\chi)^\Omega$ , namely,  $T_x S^\chi = \text{span}(d\chi_1(x), \dots, d\chi_{2s}(x))^\circ = ((V_x^\chi)^b)^\circ = (V_x^\chi)^\Omega$ . From this, using that  $S^\chi$  is symplectic one obtains  $T_x S^\chi \oplus V_x^\chi = T_x P$ . To prove that  $S \subseteq S^\chi$  is a first class constraint submanifold defined by first class constraints  $\psi_i|_{S^\chi}$ ,  $i = 1, \dots, r - 2s$ , on  $S^\chi$ , we observe first that it is immediate that  $\psi_i|_{S^\chi} = 0$ ,  $i = 1, \dots, r - 2s$ , define  $S$  regularly. It remains to show that  $X_{\psi_i|_{S^\chi}}(x) \in T_x S$ ,  $i = 1, \dots, r - 2s$ ,  $x \in S$ , where  $X_{\psi_i|_{S^\chi}}$  is the Hamiltonian vector field associated to the function  $\psi_i|_{S^\chi}$  with respect to the symplectic form  $\omega^\chi$ . This is equivalent to showing that

$$X_{\psi_i|_{S^\chi}}(x) (\psi_j|_{S^\chi}) = 0,$$

for  $x \in S$  or, equivalently,

$$\omega^\chi(x) (X_{\psi_i|_{S^\chi}}(x), X_{\psi_j|_{S^\chi}}(x)) = 0,$$

for  $x \in S$ . We know that  $\omega^\chi$  is the pullback of  $\Omega$  to  $S^\chi$  and  $\psi_i|_{S^\chi}$  is the pullback of  $\psi_i$ , via the inclusion  $S^\chi \subseteq U$ , then we have

$$\omega^\chi(x) (X_{\psi_i|_{S^\chi}}(x), X_{\psi_j|_{S^\chi}}(x)) = \Omega(x) (X_{\psi_i}(x), X_{\psi_j}(x)) = 0,$$

for  $x \in S$ , since  $\psi_i$  are first class constraints,  $i = 1, \dots, r - 2s$ . Finally, using the definitions we can easily see that  $\dim S^\chi = 2n - 2s$  and that  $\dim S = 2n - r$  and from (a) we know that  $\dim \ker \omega = r - 2s$ . We can conclude that  $2n - 2s = \dim S + \dim \ker \omega$ .

(d) We know that for  $x \in S$ ,  $\dim T_x S = 2n - r$ , and  $\dim \ker \omega(x) = r - 2s$ ; then using (3.12) we obtain  $\dim V_x = 2s$ . Also from (3.12) we immediately deduce by applying  $\Omega$  to both sides,

$$V^\Omega \cap (\ker \omega)^\Omega = TS,$$

in particular  $TS \subseteq V^\Omega$ . Let  $g$  be a given Riemannian metric on  $P$  and let  $W_x$  be the  $g$ -orthogonal complement of  $T_x S$  in  $V_x^\Omega$ , in particular,  $W_x \oplus T_x S = V_x^\Omega$ , for each  $x \in S$ . Define

$$S^V = \{\exp(tw_x) \mid w_x \in W_x, g(x)(w_x, w_x) = 1, |t| < \tau(x), x \in S\}.$$

By choosing  $\tau(x)$  appropriately one can ensure that  $S^V$  is a submanifold and, moreover, it is easy to see from the definition of  $S^V$  that  $T_x S^V = W_x \oplus T_x S = V_x^\Omega$ , for each  $x \in S$ . We leave for later the proof that  $S^V \cap U$  is a symplectic submanifold of  $P$ , for an appropriate choice of the open set  $U$ , which amounts to choosing  $\tau(x)$  appropriately.

Assume that  $d\chi'_1(x), \dots, d\chi'_{2s}(x)$  are linearly independent for  $x \in S^V \cap U'$ . Since  $\langle d\chi'_i(x), V_x^\Omega \rangle = \langle d\chi'_i(x), T_x S^V \rangle = 0$  for  $x \in S$  and  $i = 1, \dots, 2s$ , we can deduce that  $d\chi'_i(x) \in (V_x^\Omega)^\circ$ , that is,  $d\chi'_i(x) \in V_x^b$ . Then using (3.13), we see that  $d\chi'_1(x), \dots, d\chi'_{2s}(x), d\psi_1(x), \dots, d\psi_{r-2s}(x)$  are linearly independent and span  $V^b \oplus (\ker \omega)^b = (TS)^\circ$ . If  $\det(\{\chi'_i, \chi'_j\}(x)) = 0$  for some  $x \in S$  then  $\lambda^i \{\chi'_i, \chi'_j\}(x) = 0$ , where at least some  $\lambda^i \neq 0$ ,  $i = 1, \dots, 2s$ . Let  $\lambda^i \chi'_i = \varphi$ , then  $\{\varphi, \chi'_j\}(x) = 0$ ,  $j = 1, \dots, 2s$ . On the other hand, since  $\varphi|_S = 0$ , then  $\{\varphi, \psi_j\}(x) = 0$ ,  $j = 1, \dots, r - 2s$ . We can conclude that  $\varphi \in I^{(S, U')}$  and then  $X_\varphi(x) \in \ker \omega(x)$ , in particular,  $X_\varphi(x) = \mu^j X_{\psi_j}(x)$ , which implies  $\lambda^i d\chi'_i(x) = d\varphi(x) = \mu^j d\psi_j(x)$ , contradicting the linear independence of  $d\chi'_1(x), \dots, d\chi'_{2s}(x), d\psi_1(x), \dots, d\psi_{r-2s}(x)$ .

It follows from which precedes that by replacing  $\chi_i$  by  $\chi'_i$ ,  $i = 1, \dots, 2s$  and  $S$  by  $S \cap U'$  all the properties stated in (a), (b) and (c) are satisfied. In particular,  $S \cap U' = S^{\chi'}$  and  $S \cap U'$  is symplectic. It is now clear that by covering  $S$  with open subsets like the  $U'$  we can define  $U$  as being the union of all such open subsets and one obtains that  $S \cap U$  is a symplectic submanifold.

(e) Let  $x \in S^\chi$ . Then, since  $F_\chi = F - \chi_i c_\chi^{ij} \{\chi_j, F\}$ , we obtain

$$\{F_\chi, \chi_k\}(x) = \{F, \chi_k\}(x) - \{\chi_i, \chi_k\}(x) c_\chi^{ij} \{\chi_j, F\}(x) = \{F, \chi_k\}(x) + \{\chi_k, F\}(x) = 0.$$

Using this we obtain

$$\begin{aligned} \{F_\chi, G_\chi\}(x) &= \{F_\chi, G - \chi_k c_\chi^{kl} \{\chi_l, G\}\}(x) = \{F_\chi, G\}(x) \\ &= \{F, G\}(x) - \{\chi_i, G\} c_\chi^{ij} \{\chi_j, F\} = \{F, G\}_{(\chi)}(x). \end{aligned}$$

For any  $F \in \mathcal{F}(P)$ ,  $x \in S^\chi$  and  $k = 1, \dots, 2s$ , we have  $X_{F_\chi}(x) \chi_k = \{\chi_k, F_\chi\}(x) = 0$ , so  $X_{F_\chi}(x) \in T_x S^\chi$ . Therefore, for any  $Y_x \in T_x S^\chi$ ,

$$\begin{aligned} \omega^\chi(x) (X_{F_\chi}(x), Y_x) &= \Omega(x) (X_{F_\chi}(x), Y_x) = dF_\chi(x) Y_x = d(F_\chi|_{S^\chi})(x) Y_x \\ &= d(F|_{S^\chi})(x) Y_x = \omega^\chi(x) (X_{F|_{S^\chi}}(x), Y_x), \end{aligned}$$

which shows that  $X_{F_\chi}(x) = X_{F|S^\chi}(x)$ , where both Hamiltonian vector fields are calculated with the symplectic form  $\omega^\chi$ . Using this, for any  $G \in \mathcal{F}(P)$  and any  $x \in S^\chi$ , one obtains

$$\{G_\chi, F_\chi\}(x) = X_{F_\chi}(x)G_\chi = X_{F_\chi}(x)G = X_{F|S^\chi}(x)G|S^\chi = \{G|S^\chi, F|S^\chi\}_\chi(x).$$

The equality  $X_{(\chi),F}(x) = X_{F_\chi}(x) = X_{F|S^\chi}(x)$  is an immediate consequence of the previous ones. The Jacobi identity for the bracket  $\{, \}_\chi$  follows using the previous formulas, namely, for  $x \in S^\chi$ , one obtains

$$\{\{F, G\}_{(\chi)}, H\}_{(\chi)}(x) = \{\{F, G\}_{(\chi)}|S^\chi, H|S^\chi\}_\chi(x) = \{\{F|S^\chi, G|S^\chi\}_\chi, H|S^\chi\}_\chi(x),$$

where the bracket in the last term is the canonical bracket on the symplectic manifold  $S^\chi$ , for which the Jacobi identity is well known to be satisfied.

(f) The equalities (3.14) and (3.15) are proven in a straightforward way. The equations (3.16) and (3.17) follow easily using a technique similar to the one used in (e). Using all this, the proof of the Jacobi identity for the bracket  $\{, \}_\chi$  on  $U$  goes as follows. Let  $x \in U$  and let  $C$  be such that  $x \in S^{\chi^C}$ . For  $F, G, H \in \mathcal{F}(U)$  using (e) we know that the Jacobi identity holds for  $\{, \}_{(\chi^C)}$  on  $S^{\chi^C}$ . But then, according to (3.15) it also holds for  $\{, \}_\chi$  for all  $x \in S^{\chi^C}$ . Now we will prove that  $S^{\chi^C}$  are the symplectic leaves. Since they are defined by equations  $\chi_i^C = 0$ ,  $i = 1, \dots, 2s$  on  $U$  we need to prove that  $\{F, \chi_i^C\}_\chi(x) = 0$ ,  $x \in S^C$ , for all  $F \in \mathcal{F}(U)$ ,  $i = 1, \dots, 2s$ . Using (3.15) and (3.16) we see that  $\{F, \chi_i^C\}_{(\chi)}(x) = \{F, \chi_i^C\}_{(\chi^C)}(x) = \{F|S^{\chi^C}, \chi_i^C|S^{\chi^C}\}_{\chi^C}(x) = 0$ . To finish the proof, observe first that, since  $\chi_i = 0$ ,  $\psi_i = 0$ ,  $i = 1, \dots, 2s$ ,  $j = 1, \dots, r - 2s$  define regularly the submanifold  $S \subseteq U$ , by shrinking  $U$  if necessary and for all  $C$  sufficiently small, we have that  $\chi_i^C = 0$ ,  $\psi_i = 0$ ,  $i = 1, \dots, 2s$ ,  $j = 1, \dots, r - 2s$  define regularly a submanifold  $S^C \subseteq U$  and therefore  $\psi_j|S^{\chi^C} = 0$ ,  $j = 1, \dots, r - 2s$ , define regularly  $S^C$  as a submanifold of  $S^{\chi^C}$ . To prove that it is a first class constraint submanifold and that  $\psi_j|S^{\chi^C}$ ,  $j = 1, \dots, r - 2s$ , are first class constraints, that is  $\psi_j|S^{\chi^C} \in I^{(S^C, S^{\chi^C})}$ ,  $j = 1, \dots, r - 2s$ , we proceed in a similar fashion as we did in (c), replacing  $\chi$  by  $\chi^C$ . The fact that  $\psi_j|S^{\chi^C}$ ,  $j = 1, \dots, r - 2s$ , are first class constraints defining  $S^C$  implies that  $\dim S^{\chi^C} = \dim S^C + \dim \ker \omega^C$ . From the definitions one can deduce that  $\dim S^C = \dim S$  and  $\dim S^{\chi^C} = \dim S^\chi$ , therefore  $\dim \ker \omega = \dim \ker \omega^C$ . ■

The following theorem summarizes the essential part of the previous lemma.

**Theorem 3.24.** *Let  $(P, \Omega)$  be a symplectic manifold,  $S \subseteq P$  and let  $\omega$  be the pullback of  $\Omega$  to  $S$ . Assume that  $\ker \omega$  has constant rank. Let  $V$  be a vector subbundle of  $TP|S$  such that  $V \oplus \ker \omega = (TS)^\Omega$ . Then there is a symplectic submanifold  $S^V$  containing  $S$  of dimension  $\dim S + \dim \ker \omega$  such that the condition  $T_x S^V = V_x^\Omega$ , for all  $x \in S$  holds. The vector bundle  $V^\Omega$  is called the **second class subbundle** tangent to the second class submanifold  $S^V$ . The submanifold  $S^V$  is locally defined by certain equations  $\chi_1 = 0, \dots, \chi_{2s} = 0$  as in Lemma 3.23. Under the hypotheses of that lemma, namely, that  $2s$  is constant on a neighborhood  $U$  of  $S$ ,  $S^V$  will be one of the symplectic leaves of the Dirac bracket  $\{, \}_\chi$ .*

*Remark 3.25.* (a) Under our strong regularity conditions the symplectic leaves of the Dirac bracket give a (local) regular foliation of a neighborhood of the final constraint submanifold  $S$ . This implies by the Weinstein splitting theorem (Weinstein [1983]) that there are local charts where the Dirac bracket is constant.

(b) The tangent second class subbundle  $V^\Omega$  in a sense (modulo tangency) classifies all the possible second class constraint submanifolds containing a given submanifold  $S \subseteq P$ . It carries enough information to write the Dirac brackets along the final constraint submanifold  $S$  and therefore also equations of motion, as we show in section 3.4.

### 3.4 Equations of motion

We are going to describe equations of motion in the abstract setting of section 3.3, that is, a symplectic manifold  $(P, \Omega)$  and a submanifold  $S \subseteq P$ , defined regularly by equations  $\phi_i = 0$ ,  $i = 1, \dots, a$ . For this purpose, we need to introduce in this abstract setting, by definition, the notions of *primary constraints*, *primary constraint submanifold* and the *energy*.

The **primary constraint submanifold** is a given submanifold  $S' \subseteq P$  containing  $S$ , and in this context,  $S$  will be called the final constraint. We will assume without loss of generality that  $S'$  is defined regularly by the equations  $\phi_i = 0$ ,  $i = 1, \dots, a'$ , with  $a' \leq a$ , where each  $\phi_i$ ,  $i = 1, \dots, a'$  will be called a **primary constraint** while each  $\phi_i$ ,  $i = a' + 1, \dots, a$  will be called a **secondary constraint**, for obvious reasons. In this abstract setting the **energy** is by definition a given function  $\mathcal{E} \in \mathcal{F}(P)$ .

The equations of motion can be written in the Gotay-Nester form,

$$\Omega(x)(\dot{x}, \delta x) = d\mathcal{E}(x)(\delta x), \quad (3.18)$$

where  $(x, \dot{x}) \in T_x S$ , for all  $\delta x \in T_x S'$ . We require that for each  $x \in S$ , the space of solutions  $(x, \dot{x})$  of (3.18) is nonempty and forms an affine bundle.

Now we will transform this equation into an equivalent Poisson equation using the Dirac bracket.

The condition  $\{\mathcal{E}, \psi\}(x) = 0$ , for all  $x \in S$  and all first class constraints  $\psi$  will appear later as a necessary condition for existence of solutions for any given initial condition in  $S$ , so we will assume it from now on.

The **total energy** is defined by

$$\mathcal{E}_T = \mathcal{E} + \lambda^i \phi_i, \quad i = 1, \dots, a' \quad (3.19)$$

where the functions  $\lambda^i \in C^\infty(P)$ ,  $i = 1, \dots, a'$  must satisfy, by definition,  $\{\mathcal{E}_T, \phi_j\}(x) = 0$ ,  $j = 1, \dots, a$ ,  $x \in S$  or, equivalently,

$$\{\mathcal{E}, \phi_j\}(x) + \lambda^i \{\phi_i, \phi_j\}(x) = 0, \quad x \in S, \quad i = 1, \dots, a', \quad j = 1, \dots, a,$$

sum over  $i = 1, \dots, a'$ .

**Assumption 3.26.** We assume that the solutions  $(\lambda^1, \dots, \lambda^{a'})$  form a nonempty affine bundle  $\Lambda \rightarrow S$ . Equivalently, if we call  $d_\Lambda$  the dimension of the fibers,  $s' := \text{rank}(\{\phi_i, \phi_j\}(x)) = a' - d_\Lambda$  is constant on  $S$ .

By taking a sufficiently small neighborhood  $U$ , we can assume that  $\Lambda$  is a trivial bundle over  $S \cap U$ , which we extend arbitrarily as a trivial affine bundle on  $U$ .

For each section of  $\Lambda$  one has a Hamiltonian  $\mathcal{E}_T(x) = \mathcal{E}(x) + \lambda^i(x)\phi_i(x)$ ,  $x \in U$ , and the equation of motion on  $S \cap U$ ,

$$X_{\mathcal{E}_T} = (d\mathcal{E}_T)^\sharp.$$

**Equations of motion in terms of the Dirac bracket.** Local equations of motion in terms of a Dirac bracket become simpler when the second class constraints are adapted to the primary constraint, as we will explain next. This has been considered by Dirac. The case of general second class constraints is important for the global case, and will not be studied here.

Assume that the primary and final constraints  $S' \supseteq S$  are defined by

$$\phi_i = 0, i = 1, \dots, a', \text{ defines } S' \quad (3.20)$$

$$\phi_i = 0, i = 1, \dots, a', a' + 1, \dots, a, \text{ defines } S, \quad (3.21)$$

regularly on a certain neighborhood  $U$  of  $S'$ . For choosing first class and second class constraints one can proceed as follows.

**Assumption 3.27.** The rank of the matrix  $\Phi(x) = (\{\phi_i, \phi_j\}(x))$ ,  $i, j = 1, \dots, a$ , which we will call  $2s$ , is constant on an open neighborhood  $U$  of  $S'$ .

By shrinking  $U$  if necessary, one can choose  $\{\chi_1, \dots, \chi_{2s}\} \subseteq \{\phi_1, \dots, \phi_a\}$  in such a way that  $(\{\chi_i, \chi_j\}(x))$  is a regular submatrix. Without loss of generality, assume that these functions are the last  $2s$ , that is,  $\chi_i = \phi_{a-2s+i}$ ,  $i = 1, \dots, 2s$ . For each  $\phi_i$ ,  $i = 1, \dots, a - 2s$ , there are  $\alpha_i^k(x)$  such that  $\psi_i(x) := \phi_i(x) + \alpha_i^k(x)\chi_k(x)$ ,  $k = 1, \dots, 2s$ , defined on  $U$  satisfy

$$\{\psi_i, \phi_l\}(x) = \{\phi_i, \phi_l\}(x) + \alpha_i^k(x)\{\chi_k, \phi_l\}(x) = 0,$$

for all  $x \in S$ , and all  $l = 1, \dots, a$ . Therefore, these  $\psi_i$  are first class constraints and  $\{\chi_1, \dots, \chi_{2s}\}$  are second class constraints, and for each  $x \in U$  the

$$d\psi_1(x), \dots, d\psi_{a-2s}(x), d\chi_1(x), \dots, d\chi_{2s}(x),$$

are linearly independent,  $x \in U$ . We will sometimes call  $(\psi_1, \dots, \psi_{a-2s}, \chi_1, \dots, \chi_{2s})$  a complete set of first class and second class constraints for  $S$ .

Now we will show how to choose a complete set of first class and second class constraints **adapted to the primary constraints**. This, as a definition, means choosing

$$\psi_1(x), \dots, \psi_{a-2s}(x), \chi_1(x), \dots, \chi_{2s}(x),$$

as before in such a way that, for  $s'$  equal to the rank of the submatrix  $\Phi' = (\{\phi_i, \phi_j\}(x)), i = 1, \dots, a', j = 1, \dots, a$ ,

$$\psi_1, \dots, \psi_{a'-s'}, \chi_1, \dots, \chi_{s'}$$

are primary constraints.

To achieve this, let  $\{\chi'_1, \dots, \chi'_{s'}\} \subseteq \{\phi_1, \dots, \phi_{a'}\}$  such that the rows  $\{\chi'_k, \phi_j\}(x), k = 1, \dots, s',$  are linearly independent. Without loss of generality, assume that these functions are the last  $s'$ , that is,  $\chi'_i = \phi_{a'-s'+i}, i = 1, \dots, s'.$  Then, for each  $i = 1, \dots, a' - s',$  there are  $\alpha_i^k(x), k = 1, \dots, s'$  such that

$$\psi'_i(x) := \phi_i(x) + \alpha_i^k(x)\chi'_k(x)$$

defined on  $U$  satisfy

$$\{\psi'_i, \phi_l\}(x) = \{\phi_i, \phi_l\}(x) + \alpha_i^k(x)\{\chi'_k, \phi_l\}(x) = 0,$$

for all  $i = 1, \dots, a' - s', x \in S,$  and all  $l = 1, \dots, a.$  Therefore, these  $\psi'_i$  form a set of  $a' - s'$  first class constraints. Note that  $(\psi'_1, \dots, \psi'_{a'-s'}, \chi'_1, \dots, \chi'_{s'})$  is a complete set of primary constraints, in the sense that their differentials are linearly independent on  $U,$  and every primary first class constraint  $\psi$  can be written as  $\psi(x) = b_1(x)\psi'_1(x) + \dots + b_{a'-s'}(x)\psi'_{a'-s'}(x).$

In order to obtain a complete set of first class and second class constraints we first choose  $\{\chi''_{s'+1}, \dots, \chi''_{2s}\} \subseteq \{\phi_{a'+1}, \dots, \phi_a\}$  such that the rows  $\{\chi'_{k'}, \phi_j\}(x), \{\chi''_{k''}, \phi_j\}(x) k' = 1, \dots, s', k'' = s' + 1, \dots, 2s$  are linearly independent. This can be done because the matrix  $(\{\phi_i, \phi_j\}(x)), i, j = 1, \dots, a,$  has rank  $2s.$  As before, assume that  $\chi''_i = \phi_{a-2s+i}, i = s' + 1, \dots, 2s.$

Then, for each  $i = a' - s' + 1, \dots, a - 2s$  there are  $\alpha_i^{k'}(x), \alpha_i^{k''}(x), k' = 1, \dots, s', k'' = s' + 1, \dots, 2s,$  such that  $\psi''_i(x) := \phi_{s'+i}(x) + \alpha_i^{k'}(x)\chi'_{k'}(x) + \alpha_i^{k''}(x)\chi''_{k''}(x), k' = 1, \dots, s', k'' = s' + 1, \dots, 2s,$  satisfies

$$\{\psi''_i, \phi_l\}(x) = \{\phi_{s'+i}, \phi_l\}(x) + \alpha_i^{k'}(x)\{\chi'_{k'}, \phi_l\}(x) + \alpha_i^{k''}(x)\{\chi''_{k''}, \phi_l\}(x) = 0,$$

for all  $x \in S,$  and all  $l = 1, \dots, a.$

It can be shown using the definitions that

$$(d\psi'_1(x), \dots, d\psi'_{a'-s'}(x), d\psi''_{a'-s'+1}(x), \dots, d\psi''_{a-2s}(x), \\ d\chi'_1(x), \dots, d\chi'_{s'}(x), d\chi''_{s'+1}(x), \dots, d\chi''_{2s}(x))$$

are linearly independent for  $x \in S.$  Note that no nontrivial linear combination of the  $\psi''_i$  is a primary first class constraint.

Therefore,

$$(\psi_1, \dots, \psi_{a-2s}, \chi_1, \dots, \chi_{2s}) := \\ (\psi'_1, \dots, \psi'_{a'-s'}, \psi''_{a'-s'+1}, \dots, \psi''_{a-2s}, \chi'_1, \dots, \chi'_{s'}, \chi''_{s'+1}, \dots, \chi''_{2s})$$

form a complete set of first class and second class constraints adapted to the primary constraints.

We will now take advantage of these adapted constraints to write the equations of motion in a simpler way. The total energy defined on  $U$  is

$$\mathcal{E}_T = \mathcal{E} + \lambda^i \psi'_i + \mu'^j \chi'_j,$$

$i = 1, \dots, a' - s'$ ,  $j = 1, \dots, s'$ . As usual, we impose on  $\mathcal{E}_T$  the condition that it is first class, which means that  $\{\mathcal{E}, \psi_i\}(x) = 0$ ,  $i = 1, \dots, a - 2s$  and

$$\{\mathcal{E}, \chi_i\}(x) + \mu'^j \{\chi'_j, \chi_i\}(x) = 0,$$

$i = 1, \dots, 2s$ ,  $x \in S$ . From this we obtain well-defined  $\mu'^j(x)$ ,  $x \in S$ , which we extend arbitrarily for  $x \in U$ . Then the equations of motion are written as

$$X_{\mathcal{E} + \mu'^j \chi'_j}(x) + \lambda^i X_{\psi'_i}(x),$$

$x \in S$ , where  $\lambda^i$ ,  $i = 1, \dots, a' - s'$ , are completely arbitrary real numbers. We can also write this equation in the Poisson form

$$\dot{F} = \{F, \mathcal{E} + \mu'^j \chi'_j\} + \lambda^i \{F, \psi'_i\}.$$

Now we will see a simpler expression for the equations using Dirac brackets rather than the canonical bracket. Since  $\mathcal{E}_T$  is first class, then  $X_{\mathcal{E}_T}(x) = X_{(\chi), \mathcal{E}_T}(x)$  for  $x \in S$  (see Lemma 3.23(e)). Then the equations of motion can be written as

$$\dot{F} = \{F, \mathcal{E}_T\}_{(\chi)}. \quad (3.22)$$

This equation, obtained in an abstract setting, is similar to the one in page 42 of [Dirac \[1964\]](#) for the case  $P = T^*Q$ .

We can rewrite the total energy as

$$\mathcal{E}_T = \mathcal{E} + \lambda^i \psi'_i + \mu'^j \chi'_j = \mathcal{E} + \lambda^i (\phi_i + \alpha_i^k \chi'_k) + \mu'^j \chi'_j,$$

where  $i = 1, \dots, a' - s'$  and  $j = 1, \dots, s'$ , and therefore (3.22) can be written as

$$\dot{F} = \{F, \mathcal{E}\}_{(\chi)} + \lambda^i \{F, \phi_i\}_{(\chi)},$$

or equivalently,

$$X_{(\chi), \mathcal{E}_T}(x) = X_{(\chi), \mathcal{E}}(x) + \lambda^i X_{(\chi), \phi_i}(x),$$

$x \in S$ .

We will denote

$$\mathcal{E}_{AT} = \mathcal{E} + \lambda^i \phi_i \quad (3.23)$$

$i = 1, \dots, a' - s'$ , and call it the **abridged total energy**. It has the property that it provides the same dynamical information as the total energy in terms of the Dirac bracket. Moreover, for each  $x \in S$  the map  $\mathbb{R}^{a'-s'} \rightarrow \Lambda_x$  given by  $(\lambda^1, \dots, \lambda^{a'-s'}) \mapsto X_{(\chi), \mathcal{E}_{AT}}(x)$  is an isomorphism. In particular, for any solution curve  $x(t)$  of  $X_{\mathcal{E}_T}$  on  $S$ , we have

$$(x(t), \dot{x}(t)) = X_{(\chi), \mathcal{E}_{AT}}(x(t))$$

for a (uniquely determined) choice of  $\lambda^i(t)$ ,  $i = 1, \dots, a' - s'$ .

We have proven the following theorem.



**Theorem 3.28.** *Let the primary  $S'$  and final constraint  $S$  be submanifolds defined regularly by equations (3.20), (3.21) on a neighborhood  $U \subseteq P$  of  $S'$ , where  $(P, \Omega)$  is a symplectic manifold. Let  $2s$  be the rank of the matrix  $\Phi = (\{\phi_i, \phi_j\}(x))$ ,  $i, j = 1, \dots, a$ , which, by Assumption 3.27, is constant on  $U$ . Let  $s'$  be the rank of the matrix  $\Phi' = (\{\phi_i, \phi_j\}(x))$ ,  $i = 1, \dots, a', j = 1, \dots, a$ , which, by Assumption 3.26 is constant on  $S$ . Choose  $s'$  linearly independent rows  $(\{\chi'_i, \phi_j\}(x))$ ,  $i = 1, \dots, s', j = 1, \dots, a$  of  $\Phi'$ , say w.l.o.g.  $\chi'_i = \phi_{a'-s'+i}$ ,  $i = 1, \dots, s'$ . Then for each  $i = 1, \dots, a'-s'$  one can find coefficients  $\alpha_i^k(x)$  such that  $\psi'_i(x) := \phi_i(x) + \alpha_i^k(x)\chi'_k(x)$ ,  $k = 1, \dots, s'$ , defined on  $U$  constitute a maximal set of linearly independent primary first class constraints. Then  $(\psi'_1, \dots, \psi'_{a'-s'}, \chi'_1, \dots, \chi'_{s'})$  is a complete set of primary constraints and one can extend it to a complete set of first class and second class constraints*

$$(\psi_1, \dots, \psi_{a-2s}, \chi_1, \dots, \chi_{2s}) := (\psi'_1, \dots, \psi'_{a'-s'}, \psi''_{a'-s'+1}, \dots, \psi''_{a-2s}, \chi'_1, \dots, \chi'_{s'}, \chi''_{s'+1}, \dots, \chi''_{2s}).$$

Let  $\mathcal{E}: U \rightarrow \mathbb{R}$  be a given energy satisfying  $\{\mathcal{E}, \psi_i\}(x) = 0$ ,  $x \in S$ ,  $i = 1, \dots, a - 2s$ . Consider the following abridged total energy

$$\mathcal{E}_{AT} = \mathcal{E} + \lambda^i \phi_i,$$

$i = 1, \dots, a' - s'$ , where  $\lambda^i_i$  are arbitrary real parameters. The corresponding Hamiltonian vector field with respect to the Dirac bracket is

$$X_{(\chi), \mathcal{E}_T}(x) = X_{(\chi), \mathcal{E}}(x) + \lambda^i X_{(\chi), \phi_i}(x), \quad (3.24)$$

$x \in U$ . For  $x \in S$ , this vector field is tangent to  $S$  and gives the equations of motion. Equivalently, we can write the equations of motion as

$$\dot{F} = \{F, \mathcal{E}\}_{(\chi)} + \lambda^i \{F, \phi_i\}_{(\chi)}. \quad (3.25)$$

Since  $X_{(\chi), \mathcal{E}_T}$  is tangent to  $S$ , the evolution of a function  $f$  on  $S$  is given by (3.25) for any  $F$  such that  $f = F|_S$ .

*Remark 3.29.* Equation (3.25) represents an elaboration on the equation (3.22). It has interesting advantages which may be useful in practice, because it only requires calculating  $(\chi'_1, \dots, \chi'_{s'}, \chi''_{s'+1}, \dots, \chi''_{2s})$ , which is a set of second class constraints adapted to the primary constraint. It will be generalized to extend its applicability to examples where the constraints are *foliated*, as we will see in section 7.

*Remark 3.30.* The geometric meaning of  $s'$  is given by the fact that  $a' - s'$  is the number of independent primary first class constraints  $\psi'_i$ ,  $i = 1, \dots, a' - s'$ . This number coincides with  $d^{(c)}(x)$  and with the dimension of the leaves of the foliation  $\tilde{K}_c$  of  $M_c$  (see Theorem 3.19). Also, the Hamiltonian vector fields  $X_{\psi'_i}, \dots, X_{\psi'_{a'-s'}}$  are tangent to the leaves of  $\tilde{K}_c$ . Using the formula  $[X_f, X_g] = X_{-\{f, g\}}$  it is easy to prove that the iterated Poisson brackets of the  $\psi'_i$  give functions whose Hamiltonian vector fields are tangent to the leaves of  $\tilde{K}_c$ . Each one of these functions, let us call it  $f$ , satisfies that  $X_f(x) = \lambda^i(x)X_{\psi'_i}(x)$  for all  $x \in M_c$ , but not necessarily  $f$

is a linear combination of the  $\psi'_i$ , on a neighborhood of  $M_c$ , so it might not be a primary first-class constraint. One can prove that such a function is zero on  $M_c$ , but since  $df(x) = \lambda^i(x)d\psi'_i(x)$  for all  $x \in M_c$ , it can not be included in a set of constraint functions defining  $M_c$  regularly and containing the  $\psi'_i$ , in other words, it is not an *independent* secondary first class constraint. There are examples in the literature where the Poisson bracket of two primary first class constraints is an independent secondary first class constraint, for instance, that is the case in Henneaux and Teitelboim [1992, p. 2]; for that example Assumption 3.18 (a) does not hold (but 3.18 (b) does).

According to Dirac [1964], pages 23–24, primary first class constraints and their iterated brackets represent transformations that do not change the physical state of the system, which in this case is implied directly from the requirement of determinism. We have seen that under our regularity conditions, those iterated brackets do not really add any new such transformations besides the ones given by the primary first class constraints.

In addition, note that  $a - 2s$  is also the dimension of the kernel of the presymplectic form on  $S$  (see Theorem 3.23 (a)), which is the number of first class constraints, and also the dimension of the leaves of  $K_c$  (see Assumption 3.8). Since the Hamiltonian vector fields  $X_\psi$  corresponding to the first class constraints generate the integrable distribution associated to  $K_c$ , then the transformations that preserve the leaves of  $K_c$  are those which preserve the physical state. So, points of  $\bar{M}_c = M_c/K_c$  would represent exactly the physical states.

Dirac [1964, p. 25] introduces the notion of extended Hamiltonian. In our context, we should define the notion of *extended energy*,

$$\mathcal{E}_E = \mathcal{E}_T + \lambda^{''i}\psi''_i, \quad (3.26)$$

$i = 1, \dots, a - 2s - (a' - s')$  where the  $\psi^{''i}$  represent the secondary first class constraints. Both  $\mathcal{E}_E$  and  $\mathcal{E}_T$  give the same dynamics on the quotient manifold  $\bar{M}_c$ . If  $(M_c, \omega_c)$  is symplectic, there are only second class constraints among the  $\phi_i$ ,  $i = 1, \dots, a$ , and conversely. In this case,  $\mathcal{E}_E = \mathcal{E}_T$ .

## 4 Dirac structures

The rest of the paper is devoted to generalizing the Dirac and Gotay-Nester theories of constraints. In this section we review a few basic facts related to Dirac structures and Dirac dynamical systems following Courant [1990] and Yoshimura and Marsden [2006a,b], which are of direct interest for the present paper.

We introduce the flat  $D^\flat$  and the orthogonal  $D$  operators with respect to a given Dirac structure, which is important to describe the algorithm, in the next section. In fact, we obtain a technique which imitates the Gotay-Nester technique, by replacing  $\omega$  by  $D$ .

**Dirac structures on vector spaces.** Let  $V$  be an  $n$ -dimensional vector space and let  $V^*$  be its dual space. Define the *symmetric pairing*  $\langle\langle \cdot, \cdot \rangle\rangle$  on  $V \oplus V^*$  by

$$\langle\langle (v_1, \alpha_1), (v_2, \alpha_2) \rangle\rangle = \langle \alpha_1, v_2 \rangle + \langle \alpha_2, v_1 \rangle,$$

for  $(v_1, \alpha_1), (v_2, \alpha_2) \in V \oplus V^*$  where  $\langle \cdot, \cdot \rangle$  is the natural pairing between  $V^*$  and  $V$ . A *linear Dirac structure* on  $V$ , (sometimes called simply a *Dirac structure* on  $V$ ), is a subspace  $D \subseteq V \oplus V^*$  such that  $D = D^\perp$ , where  $D^\perp$  is the orthogonal of  $D$  relative to the pairing  $\langle\langle \cdot, \cdot \rangle\rangle$ . One can easily check the following result.

**Lemma 4.1.** *A vector subspace  $D \subseteq V \oplus V^*$  is a Dirac structure on  $V$  if and only if it is maximally isotropic with respect to the symmetric pairing. A further equivalent condition is given by  $\dim D = n$  and  $\langle\langle (v_1, \alpha_1), (v_2, \alpha_2) \rangle\rangle = 0$  for all  $(v_1, \alpha_1), (v_2, \alpha_2) \in D$ .*

One of the main examples is the Dirac structure naturally defined on a presymplectic vector space  $(V, \omega)$  by

$$D_\omega = \{(v, \alpha) \in V \oplus V^* \mid \alpha = \omega^\flat(v)\}.$$

Here the definition of  $\omega^\flat$  is the standard one, namely,  $\omega^\flat: V \rightarrow V^*$  is defined by  $\omega^\flat(v)(w) = \omega(v, w)$ , for all  $v, w \in V$ . We also recall the definition of orthogonality on  $V$  associated to a given presymplectic form  $\omega$ . For a given subset  $W \subseteq V$  we define the  $\omega$ -orthogonal complement  $W^\omega$  by

$$W^\omega = \{v \in V \mid \omega(v, w) = 0 \text{ for all } w \in W\}.$$

We recall some standard facts in the following lemma, omitting the proof.

**Lemma 4.2.** *Let  $\omega$  be a presymplectic form on a vector space  $V$  and let  $W \subseteq V$  be any vector subspace. Then  $W^\flat := \omega^\flat(W) = (W^\omega)^\circ$ , where the right hand side denotes the annihilator of  $W^\omega$ .*

The following proposition is a direct consequence of propositions 1.1.4 and 1.1.5 in Courant [1990].

**Proposition 4.3.** *Given a Dirac structure  $D \subseteq V \oplus V^*$  define the subspace  $E_D \subseteq V$  to be the projection of  $D$  on  $V$ . Also, define the 2-form  $\omega_D$  on  $E_D$  by  $\omega_D(v, w) = \alpha(w)$ , where  $v \oplus \alpha \in D$ . (One checks that this definition of  $\omega_D$  is independent of the choice of  $\alpha$ ). Then,  $\omega_D$  is a skew form on  $E_D$ . Conversely, given a vector space  $V$ , a subspace  $E \subseteq V$  and a skew form  $\omega$  on  $E$ , one sees that  $D_\omega = \{v \oplus \alpha \mid v \in E, \alpha(w) = \omega(v, w) \text{ for all } w \in E\}$  is the unique Dirac structure  $D$  on  $V$  such that  $E_D = E$  and  $\omega_D = \omega$ .*

*$D$  is the Dirac structure associated to a presymplectic form  $\omega$  on  $V$ , as explained before, if and only if  $E_D = V$  and  $\omega = \omega_D$ .*

**The operators  $D^\flat$  and  $D$ .** There is a natural extension, which is important in the present paper, of the previous definition of  $\omega^\flat$  for the case of a general Dirac structure  $D \subseteq V \oplus V^*$ . For a given Dirac structure  $D \subseteq V \oplus V^*$  and given  $X \in E_D$  define the set  $D^\flat(X)$ , sometimes denoted  $X^\flat$  for short, by

$$D^\flat(X) = \{\alpha \in V^* \mid (X, \alpha) := X \oplus \alpha \in D\}.$$

Note that  $D$  is the Dirac structure associated to a presymplectic form on  $V$ , that is,  $D = D_\omega$ , if and only if  $E_D = V$  and for each  $X \in V$ , the set  $D^\flat(X)$  has a single element, more precisely,  $D^\flat(X) = \{\omega^\flat(X)\}$ . In this sense,  $D^\flat(X)$  generalizes  $\omega^\flat(X)$ . For a given subset  $W \subseteq E_D$  define  $D^\flat(W)$ , also denoted  $W^\flat$ , by

$$D^\flat(W) = \bigcup_{X \in W} D^\flat(X).$$

If  $W$  is a subspace, then  $D^\flat(W)$  is a subspace of  $V^*$ .

It is straightforward to check that for all  $X, Y \in E_D$ ,

$$\{\omega_D(X, Y)\} = D^\flat(X)(Y) = -D^\flat(Y)(X) = -\{\omega_D(Y, X)\}.$$

The notion of orthogonal complement with respect to  $\omega$  can be generalized as follows. For any subset  $W \subseteq E_D$  define  $W^D$  by

$$W^D = \{X \in V \mid D^\flat(Y)(X) = \{0\}, \text{ for all } Y \in W\}.$$

Clearly,

$$W^D \cap E_D = \{X \in E_D \mid D^\flat(Y)(X) = \{0\}, \text{ for all } Y \in W\}.$$

It is easy to check that for any subspace  $W \subseteq E_D$  one has

$$W^D \cap E_D = W^{\omega_D}.$$

We recall also that, since  $\omega_D: E_D \times E_D \rightarrow \mathbb{R}$  is a presymplectic form on  $E_D$ , one has, according to Lemma 4.2 and with a self-explanatory notation,

$$(W^{\omega_D})^{\circ E_D} = W^{\flat \omega_D}.$$

The following proposition generalizes Lemma 4.2 and is one of the ingredients of the constraint algorithm described in section 5.

**Proposition 4.4.** *Let  $D \subseteq V \oplus V^*$  be a given Dirac structure and let  $W \subseteq E_D$  be a given subspace. Then  $D^\flat(W) \equiv W^\flat = (W^D)^\circ$ .*

**Proof.** We first show that  $W^\flat \subseteq (W^D)^\circ$ . Let  $\alpha \in W^\flat$ , say  $\alpha \in D^\flat(X)$  for some  $X \in W$ . Since  $D^\flat(X)(Y) = \{0\}$  for all  $Y \in W^D$  by definition of  $W^D$ , it follows that  $\alpha \in (W^D)^\circ$ . To prove the converse inclusion, we first observe that it is equivalent to prove that  $(W^\flat)^\circ \subseteq W^D$ . Let  $X \in (W^\flat)^\circ$ , then  $D^\flat(Y)(X) = 0$ , for all  $Y \in W$ , which is an immediate consequence of the definitions. But, also by definition, this implies that  $X \in W^D$ . ■

**Dirac structures on manifolds.** We will give the definition and some basic properties of Dirac manifolds, following [Courant \[1990\]](#) and [Yoshimura and Marsden \[2006a,b\]](#), using the notation of the latter.

A *Dirac structure*  $D$  on a manifold  $M$  is a subbundle of the Whitney sum  $D \subseteq TM \oplus T^*M$  such that for each  $x \in M$ ,  $D_x \subseteq T_xM \oplus T_x^*M$  is a Dirac structure on the vector space  $T_xM$ . A *Dirac manifold* is a manifold with a Dirac structure on it. From Proposition 4.3 we deduce that a Dirac structure  $D$  on  $M$  yields a distribution  $E_{Dx} \subseteq T_xM$  whose dimension is not necessarily constant, carrying a presymplectic form  $\omega_D(x): E_{Dx} \times E_{Dx} \rightarrow \mathbb{R}$ , for all  $x \in M$ . We can also deduce the following theorem, whose detailed proof appears in [Yoshimura and Marsden \[2006a\]](#).

**Theorem 4.5.** *Let  $M$  be a manifold and let  $\omega$  be a 2-form on  $M$ . Given a distribution  $E$  on  $M$ , define the skew-symmetric bilinear form  $\omega_E$  on  $E$  by restricting  $\omega$  to  $E \times E$ . For each  $x \in M$  let*

$$D_{\omega_E} = \{(v_x, \alpha_x) \in T_xM \oplus T_x^*M \mid v_x \in E(x) \text{ and } \alpha_x(w_x) = \omega_E(x)(v_x, w_x) \text{ for all } w_x \in E(x)\}.$$

*Then  $D_{\omega_E} \subseteq TM \oplus T^*M$  is a Dirac structure on  $M$ . It is the only Dirac structure  $D$  on  $M$  satisfying  $E(x) = E_{Dx}$  and  $\omega_E(x) = \omega_D(x)$ , for all  $x \in M$ .*

Using Lemma 4.2 and Proposition 4.4 we can easily deduce the following proposition.

**Proposition 4.6.** *Let  $D$  be a Dirac structure on  $M$  and let  $W_x$  be a subspace of  $E_{Dx}$  for each  $x \in M$ ; then, with a self-explanatory notation, the following equalities hold for each  $x \in M$ :*

$$W_x^{b\omega_D} = (W_x^{\omega_D})^{\circ E_D} \\ W_x^b = (W_x^D)^{\circ}.$$

A Dirac structure  $D$  on  $M$  is called *integrable* if the condition

$$\langle L_{X_1}\alpha_2, X_3 \rangle + \langle L_{X_2}\alpha_3, X_1 \rangle + \langle L_{X_3}\alpha_1, X_2 \rangle = 0$$

is satisfied for all pairs of vector fields and 1-forms  $(X_1, \alpha_1), (X_2, \alpha_2), (X_3, \alpha_3)$  that take values in  $D$  and where  $L_X$  denotes the Lie derivative along the vector field  $X$  on  $M$ . This definition encompasses the notion of closedness for presymplectic forms and Jacobi identity for brackets. The following fundamental theorem was proven in [Courant \[1990\]](#).

**Theorem 4.7.** *Let  $D$  be an integrable Dirac structure on a manifold  $M$ . Then the distribution  $E_D$  is involutive. If, moreover, the hypotheses of the Stefan-Sussmann theorem ([Sussmann \[1973\]](#)) are satisfied, for each  $x \in M$  there exists a uniquely determined embedded submanifold  $S$  of  $M$  such that  $x \in S$  and  $T_yS = E_{Dy}$  for all  $y \in S$ . In other words,  $S$  is an integral submanifold of  $E_D$ . Each integral submanifold  $S$  carries a presymplectic form  $\omega_{D,S}$  defined by  $\omega_{D,S}(x) = \omega_D(x)$ , for each  $x \in S$ .*

## 4.1 Dirac dynamical systems

Lagrangian and Hamiltonian mechanics has been developed in interaction with differential, symplectic, and Poisson geometry; a few references are [Abraham and Marsden \[1978\]](#); [Arnol'd \[1989\]](#); [de León and Rodrigues \[1989\]](#); [Cordero, Dodson, and de León \[1989\]](#); [Marsden and Ratiu \[1994\]](#); [Cendra, Marsden, and Ratiu \[2001b\]](#); [Holm \[2011a,b\]](#). Dirac dynamical systems in the integrable case represent a synthesis and a generalization of both.

Nonholonomic mechanics represents a generalization of Lagrangian and Hamiltonian mechanics and is a long-standing branch of mechanics, engineering and mathematics. Some references and historical accounts on the subject are [Neĭmark and Fufaev \[1972\]](#); [van der Schaft and Maschke \[1994\]](#); [Zenkov, Bloch, and Marsden \[1998\]](#); [Cendra, Marsden, and Ratiu \[2001a\]](#); [Borisov and Mamaev \[2002\]](#); [Cortés \[2002\]](#); [Bloch \[2003\]](#); [Cendra, Ibort, de León, and Martín de Diego \[2004\]](#); [Cortés, de León, Marrero, and Martínez \[2009\]](#); [Grabowski, de León, Marrero, and Martín de Diego \[2009\]](#); [Balseiro, de León, Marrero, and Martín de Diego \[2009\]](#); [Holm \[2011a,b\]](#). Some references that are more closely related to this paper are [Yoshimura and Marsden \[2006a,b\]](#).

Dirac dynamical systems (1.1) in the not necessarily integrable case may be viewed as a synthesis and a generalization of nonholonomic mechanics from the Lagrangian and the Hamiltonian points of view. They can be written equivalently as a collection of systems of the type

$$i_{\dot{x}} \omega_D(x) = d\mathcal{E}(x)|_{E_D}. \quad (4.1)$$

A related approach has been studied in [van der Schaft \[1987, 1998\]](#); [van der Schaft and Maschke \[1994\]](#); [Maschke, van der Schaft, and Breedveld \[1995\]](#); [Blankenstein and Ratiu \[2004\]](#). As it was shown in [Yoshimura and Marsden \[2006a,b\]](#), one can write nicely the equations of nonholonomic mechanics, on the Lagrangian side, using the Dirac differential. We will show how this is related to system (1.1). It was also shown in [Yoshimura and Marsden \[2006a,b\]](#) and references therein how to use Dirac structures in LC circuit theory, on the Lagrangian side. On the Hamiltonian side, Poisson brackets for LC circuits were written in [Maschke, van der Schaft, and Breedveld \[1992, 1995\]](#), see also [Maschke and van der Schaft \[1998\]](#). In [Bloch and Crouch \[1999\]](#) the Hamiltonian structure for nonlinear LC circuits in the framework of Dirac structures is investigated and simple and effective formulas are described.

In [Cendra, Marsden, Ratiu, and Yoshimura \[2011\]](#) a further unification, including reduction, is presented, which is consistent with mechanics on Lie algebroids ([Weinstein \[1996\]](#); [Cortés, de León, Marrero, Martín de Diego, and Martínez \[2006\]](#); [Martínez \[2007\]](#); [Iglesias, Marrero, Martín de Diego, Martínez, and Padrón \[2007\]](#); [Martínez \[2008\]](#); [Cortés, de León, Marrero, and Martínez \[2009\]](#)).

## 5 The constraint algorithm for Dirac dynamical systems (CAD)

Some results in this section, in particular the CAD and examples of nonholonomic mechanics, are proven for not necessarily integrable Dirac structures.

First, we shall briefly consider the case of an integrable Dirac structure. This includes the case of a constant Dirac structure on a vector space, that is, one that is invariant under translations, which includes the one used for LC circuits. Later on we will consider the general case.

**The case of an integrable Dirac manifold.** Assume that  $D$  is an integrable Dirac structure on a manifold  $M$ . Each solution curve  $x(t)$  of the system (1.1) must satisfy  $(x(t), \dot{x}(t)) \in E_D(x(t))$  for all  $t$ , which implies, according to Theorem 4.7 and equation (4.1), that it must be a solution curve to the equation

$$\omega_{D,S}(x)(\dot{x}, \cdot) = i_S^* d\mathcal{E}(x) \quad (5.1)$$

on a presymplectic leaf  $S$ , where  $i_S: S \rightarrow M$  is the inclusion, which can be solved using the Gotay-Nester algorithm. Such a procedure to solve (1.1) might be useful in those cases where the presymplectic leaves  $(S, \omega_{D,S})$  can be found easily. This occurs for instance if  $M$  is a vector space and  $D$  is a constant (i.e. translation-invariant) Dirac structure, as we will show next. However, this procedure has the drawback that in order to find a solution of (1.1) for a given initial condition, one must first find the leaf  $S$  containing that initial condition and then solve (5.1). For an initial condition on a different leaf, one has to repeat the constraint algorithm for a different corresponding equation (5.1). Because of this, even in these simple cases, working directly with the Dirac structure  $D$ , using the constraint algorithm to be developed in this section, rather than with the associated presymplectic form on a presymplectic leaf, is not only possible but also convenient, since this leads to obtaining a single equation on a final foliated constraint submanifold, as we will see.

**The case of a constant Dirac structure.** Let  $V$  be a vector space and  $\tilde{D} \subseteq V \oplus V^*$  a given linear Dirac structure. Then we have the presymplectic form  $\tilde{\omega}_{\tilde{D}}$  on  $E_{\tilde{D}}$  and the associated linear map  $\tilde{\omega}_{\tilde{D}}^b: E_{\tilde{D}} \rightarrow E_{\tilde{D}}^*$ . We consider the Dirac structure  $D \subseteq TV \oplus T^*V$  on the manifold  $V$  defined as  $D_x = (x, \tilde{D})$ , where we have used the natural identification  $TV \oplus T^*V \equiv V \times (V \oplus V^*)$ . This Dirac structure is integrable and constant, that is, invariant under translations in a natural sense, as we will show next.

For each  $x \in V$  the presymplectic leaf  $S_x$  containing  $x$  (in the sense of Theorem 4.7) is  $x + E_{\tilde{D}} \subseteq V$ . For each  $x \in V$ ,  $\tilde{\omega}_{\tilde{D}}$  induces the constant presymplectic form  $\omega_{D,S} \in \Omega^2(x + E_{\tilde{D}})$  given by

$$\omega_{D,S}(x+y)((x+y, \bar{X}), (x+y, \bar{Y})) = \tilde{\omega}_{\tilde{D}}(\bar{X}, \bar{Y}),$$

where  $y \in E_{\tilde{D}}$ , so  $x+y$  represents any point in the symplectic leaf  $x + E_{\tilde{D}}$ , or, equivalently,

$$\omega_{D,S}^b(x+y)(x+y, \bar{X})(x+y, \bar{Y}) = \tilde{\omega}_{\tilde{D}}^b(\bar{X})(\bar{Y}).$$



Consider the system

$$(z, \dot{z}) \oplus (z, d\mathcal{E}(z)) \in D. \quad (5.2)$$

This system, for a given initial condition  $z(0) = x$ , is equivalent to the following equation on the presymplectic leaf  $S$  that contains  $x$

$$\omega_{D,S}(z)(\dot{z}, \cdot) = d\mathcal{E}(z)|_{E_{Dz}}. \quad (5.3)$$

Keeping  $x$  fixed and writing  $z = x + y$ , the system (5.2) on the presymplectic leaf  $S = x + E_{\tilde{D}}$  becomes

$$\dot{y} \oplus d\mathcal{E}(x + y) \in \tilde{D}.$$

Therefore equation (5.3) is equivalent to the following equation on the subspace  $E_{\tilde{D}}$

$$\tilde{\omega}_{\tilde{D}}(\dot{y}, \cdot) = d\mathcal{E}(x + y)|_{E_{\tilde{D}}}, \quad (5.4)$$

with initial condition  $y(0) = 0$ . Equation (5.4) can be solved by the algorithm described in Gotay, Nester, and Hinds [1978] and Gotay and Nester [1979], and sketched in section 2.2. Notice that, since the presymplectic form  $\tilde{\omega}_{\tilde{D}}$  on the vector space  $E_{\tilde{D}}$  determines naturally a translation-invariant Dirac structure on the same space  $E_{\tilde{D}}$  considered as a manifold, equation (5.4) is also a Dirac dynamical system in the same way (5.2) is, but with the Dirac structure given by the presymplectic form  $\tilde{\omega}_{\tilde{D}}$  considered as a constant form on the manifold  $x + E_{\tilde{D}}$ . Because of this and also because  $\dim E_{\tilde{D}} \leq \dim V$ , equation (5.4) is, in essence, simpler than system (5.2).

**The case of a general Dirac manifold.** Now let  $D$  be a Dirac structure on  $M$  that needs not be integrable. In order to explain our algorithm for Dirac manifolds we need the following auxiliary result, involving a given subspace  $W_x$  of  $E_{Dx}$ , which is easy to prove using results from section 4.

**Lemma 5.1.** *For each  $x \in M$  we have the following equivalent conditions, where  $W_x$  is a given subspace of  $E_{Dx}$ .*

- (i) *There exists  $(x, \dot{x}) \in W_x$  such that (1.1) is satisfied.*
- (ii) *There exists  $(x, \dot{x}) \in W_x$  such that  $\omega_D(x)(\dot{x}, \cdot) = d\mathcal{E}(x)|_{E_{Dx}}$ .*
- (iii)  *$d\mathcal{E}(x)|_{E_{Dx}} \in W_x^{\flat\omega_D}$ .*
- (iv)  *$\langle d\mathcal{E}(x)|_{E_{Dx}}, W_x^{\omega_D} \rangle = \{0\}$ .*
- (v) *There exists  $(x, \dot{x}) \in W_x$  such that  $d\mathcal{E}(x) \in D^\flat(x, \dot{x})$ .*
- (vi)  *$d\mathcal{E}(x) \in D^\flat(W_x)$ .*
- (vii)  *$\langle d\mathcal{E}(x), W_x^D \rangle = \{0\}$ .*

Now we can describe the constraint algorithm for Dirac dynamical systems, called *CAD*. Following the same idea of the Gotay-Nester algorithm described in section 2.2 we should construct a sequence of constraint submanifolds.

To define the first constraint submanifold  $M_1$  we may use either one of the equivalent conditions of Lemma 5.1, with  $W_x \equiv W_{0x} = E_{Dx}$ . We want to emphasize the role of the two equivalent conditions (iv) and (vii), as they represent a formal analogy between the CAD and the Gotay-Nester algorithm. Of course, the Gotay-Nester algorithm, by definition, corresponds to the case  $E_D = TM$ .

Define

$$\begin{aligned} M_1 &= \{x \in M \mid \langle d\mathcal{E}(x)|E_{Dx}, E_{Dx}^{\omega_D} \rangle = \{0\}\} \\ &= \{x \in M \mid \langle d\mathcal{E}(x), E_{Dx}^D \rangle = \{0\}\}. \end{aligned} \quad (5.5)$$

Let us assume that  $M_1$  is a submanifold. Then we define the second constraint submanifold  $M_2 \subseteq M_1$  by either of the following equivalent conditions, in agreement with (iv) and (vii) of Lemma 5.1, with  $W_x \equiv W_{1x} = E_{Dx} \cap T_x M_1$ ,

$$\begin{aligned} M_2 &= \{x \in M_1 \mid \langle d\mathcal{E}(x)|E_{Dx}, (W_{1x})^{\omega_D} \rangle = \{0\}\} \\ &= \{x \in M_1 \mid \langle d\mathcal{E}(x), (W_{1x})^D \rangle = \{0\}\}. \end{aligned}$$

More generally we define recursively  $M_{k+1}$  for  $k = 1, 2, \dots$ , by either of the conditions

$$\begin{aligned} M_{k+1} &= \{x \in M_k \mid \langle d\mathcal{E}(x)|E_{Dx}, (W_{kx})^{\omega_D} \rangle = \{0\}\} \\ &= \{x \in M_k \mid \langle d\mathcal{E}(x), (W_{kx})^D \rangle = \{0\}\} \end{aligned} \quad (5.6)$$

with  $W_{kx} = E_{Dx} \cap T_x M_k$ . The algorithm stops and the final constraint submanifold  $M_c$  is determined by the condition  $M_{c+1} = M_c$ . The solutions  $(x, \dot{x})$  are in  $W_{cx} = E_{Dx} \cap T_x M_c$ .

We have proven that solution curves of (1.1) are exactly solution curves of

$$(x, \dot{x}) \oplus d\mathcal{E}(x) \in D_x \quad (5.7)$$

$$(x, \dot{x}) \in W_{cx}. \quad (5.8)$$

*Remark 5.2.* In some examples, it is sometimes easier to write the constraint submanifolds using condition (vi) in Lemma 5.1, that is,

$$M_{k+1} = \{x \in M_k \mid d\mathcal{E}(x) \in D^b(W_{kx})\}$$

with  $W_{kx} = E_{Dx} \cap T_x M_k$ ,  $k = 0, 1, \dots$ , where  $M_0 = M$ .

*Remark 5.3.* Formula (5.6) has a special meaning in the case of an integrable Dirac structure. In fact, let  $S$  be an integral leaf of the distribution  $E_D$ ; then by applying the Gotay-Nester algorithm, encoded in the recursion formula (2.5), to the system (5.1) one obtains a sequence of secondary constraints  $S = S_0 \supseteq S_1 \supseteq \dots \supseteq S_c$  given by the recursion formula

$$S_{k+1} = \{x \in S_k \mid \langle i_S^* d\mathcal{E}(x), (T_x S_k)^{\omega_{D,S}} \rangle = \{0\}\}. \quad (5.9)$$

But it is clear that equation (5.9) coincides with equation (5.6) since  $T_x S \equiv E_{Dx}$ ,  $T_x S_k \equiv W_{kx}$  and the presymplectic form  $\omega_{D,S}$  on  $S$  is defined by  $\omega_{D,S}(x) \equiv \omega_D(x)$  for all  $x \in S$ . We are assuming regularity conditions that ensure that the Gotay-Nester algorithm applied for each  $S$  stops after a number of steps  $c$  which does not depend on  $S$  and which is at the same time the number of steps after which the CAD stops. As a conclusion, the final constraint submanifold  $M_c$  of the CAD is foliated by leaves  $S_c$ , where  $S$  varies on the set of integral leaves of the distribution  $E_D$ . We may say that in the case of an integrable Dirac structure  $D$  the CAD is equivalent to a collection of Gotay-Nester algorithms, one for each leaf  $S$  of the distribution  $E_D$ . The final equation given by (5.7)–(5.8) becomes

$$\begin{aligned} (x, \dot{x}) \oplus d\mathcal{E}(x) &\in D_x \\ (x, \dot{x}) &\in T_x S_c, \end{aligned}$$

which is equivalent to the collection of equations

$$\begin{aligned} \omega_{D,S}(x)(\dot{x}, \cdot) &= i_S^* d\mathcal{E}(x) \\ (x, \dot{x}) &\in T_x S_c. \end{aligned}$$

In section 7 we will extend the Dirac theory of constraints. For that purpose, we will use an embedding of  $M$  in a symplectic manifold  $P$  such that the presymplectic leaves  $S$  of are presymplectic submanifolds of  $P$ . The submanifold  $M$  plays the role of a primary *foliated* constraint submanifold. The case in which there is only one leaf gives the Dirac theory.

**Solving the equation.** The constraint algorithm CAD gives a method to solve the IDE (1.1) which generalizes the Gotay-Nester method. Assume that the final constraint submanifold  $M_c$  has been determined, and consider, for each  $x \in M_c$ , the affine space

$$S_x^{(c)} := \{(x, \dot{x}) \in T_x M_c \mid (4.1) \text{ is satisfied}\},$$

which is nonempty if  $M_c$  is nonempty, a condition that will be assumed from now on. Let  $d^{(c)}(x)$  be the dimension of  $S_x^{(c)}$ .

The following theorem is one of the ingredients of our main results, and generalizes the Gotay-Nester algorithm for the case of Dirac dynamical systems (5.10) rather than Gotay-Nester systems (1.2). Its proof is not difficult, using the previous lemma, and is left to the reader.

**Theorem 5.4.** *Let  $M$  be a given manifold,  $D$  a given Dirac structure on  $M$  and  $\mathcal{E}$  a given energy function on  $M$ , and consider the Dirac dynamical system*

$$(x, \dot{x}) \oplus d\mathcal{E}(x) \in D_x. \quad (5.10)$$

*Assume that for each  $k = 1, \dots$ , the subset  $M_k$  of  $M$  defined recursively by the formulas (5.5)–(5.6) is a submanifold, called the  **$k$ -constraint submanifold**. The decreasing sequence  $M_k$  stops, say  $M_c = M_{c+1}$  (which implies  $M_c = M_{c+p}$ , for all  $p \in \mathbb{N}$ ), and call  $M_c$  the **final constraint submanifold**. Then the following hold:*

(a) For each  $x \in M_c$ , there exists  $(x, \dot{x}) \in W_{cx} = E_{Dx} \cap T_x M_c$  such that (5.10) is satisfied. The Dirac dynamical system (5.10) is equivalent to the equation

$$\omega_D(x)(\dot{x}, \cdot) = d\mathcal{E}(x)|_{E_{Dx}}, \quad (x, \dot{x}) \in W_{cx}, \quad (5.11)$$

that is, both equations have the same solution curves  $x(t) \in M_c$ .

(b) For each  $x \in M_c$ ,  $d^{(c)}(x)$  equals the dimension of  $\ker \omega_D(x) \cap T_x M_c$ .

(c) If  $d^{(c)}(x)$  is a locally constant function of  $x$  on  $M_c$  then  $S^{(c)} = \bigcup_{x \in M_c} S_x^{(c)}$  is an affine bundle with base  $M_c$ . Each section  $X$  of  $S^{(c)}$  is a vector field on  $M_c$  having the property that  $X(x) \oplus d\mathcal{E}(x) \in D_x$ , for all  $x \in M_c$ . Solution curves to such vector fields are solutions to the Dirac dynamical system (1.1). More generally, one can choose arbitrarily a time-dependent section  $X_t$ , then solution curves of  $X_t$  will be also solutions to (1.1) and those are the only solutions of (1.1). Solution curves to (1.1) are unique for any given initial condition if and only if  $d^{(c)}(x) = 0$ , for all  $x \in M_c$ .

**Local representation for  $S^{(c)}$ .** One can find a local representation for  $S^{(c)}$  by just choosing a local parametrization of  $M_c$ . Let  $x = x(z_1, \dots, z_r) \in U$ , where  $U$  is an open set and  $r$  is the dimension of  $M_c$ , be such a local parametrization. Then substitute this expression for  $x$  in (1.1) to obtain an IDE in  $z = (z_1, \dots, z_r)$ , namely

$$(x(z), D_z x(z) \cdot \dot{z}) \oplus d\mathcal{E}(x(z)) \in D_{x(z)}. \quad (5.12)$$

The local representation of  $S^{(c)}$  is given by the trivial affine bundle

$$\{(z, \dot{z}) \mid (z, \dot{z}) \text{ satisfies (5.12)}\}.$$

**Some results concerning uniqueness of solution.** Under Assumption 3.18 (b), we can prove the following lemma.

**Lemma 5.5.** (A) Existence and uniqueness of a solution curve  $x(t)$  of (5.11) for any initial condition  $x(0) \in M_c$ , and therefore also of the Dirac dynamical system (5.10), is equivalent to any of the conditions

- (i)  $\ker \omega_D(x) \cap W_{cx} = \ker \omega_D(x) \cap T_x M_c = \{0\}$ , for each  $x \in M_c$ ,
- (ii)  $W_{cx} \cap E_{Dx}^{\omega_D} = \{0\}$ , for each  $x \in M_c$ ,
- (iii)  $W_{cx} \cap E_{Dx}^{\Omega} = \{0\}$ , for each  $x \in M_c$ .

(B) If  $\omega_D|_{W_{cx}}$ , or, equivalently,  $\Omega|_{W_{cx}}$ , is symplectic for each  $x \in M_c$  then there is existence and uniqueness of solution  $x(t)$  of (5.11) for any initial condition  $x(0) \in M_c$ , and therefore also of the Dirac dynamical system (5.10). One also has that  $\omega_D|_{W_{cx}}$  is symplectic for each  $x \in M_c$  if and only if any one of the following conditions is satisfied:

- (i)  $W_{cx} \cap W_{cx}^{\omega_D} = \{0\}$ ,
- (ii)  $W_{cx} \cap W_{cx}^{\Omega} = \{0\}$ .

**Proof.** We first recall the argument from Theorem 5.4. Uniqueness of a solution  $x(t)$  of equation (5.11) for any given initial condition  $x_0 \in M_c$  (which, as we know from the CAD, must satisfy  $x(t) \in M_c$  and  $\dot{x}(t) \in W_{cx(t)}$  for all  $t$ ) holds if and only if for each  $x \in M_c$  there is a uniquely determined  $v_x \in W_{cx}$  such that

$$\omega_D(x)(v_x, \cdot) = d\mathcal{E}|_{E_{Dx}},$$

in other words, if equation (5.11) defines a vector field on  $M_c$ . In fact, if this is the case, by the general theory of ODE on manifolds we have existence and uniqueness of solution. We also recall that under Assumption 3.18 (b), the equation (5.11) defines a family of vector fields on  $M_c$ , which defines an affine distribution of constant rank whose space at the point  $x \in M_c$  is the affine space of all solutions  $v_x$  as indicated above. Using this we can deduce that uniqueness of solution  $x(t)$  of (5.11) for any initial condition  $x(0) \in M_c$ , is equivalent to the affine distribution defined above having dimension 0.

Now we shall prove the equivalence between uniqueness of solution and (A). Let  $v_x \in W_{cx}$  satisfying equation (5.11), that is,

$$\omega_D(x)(v_x, \cdot) = d\mathcal{E}|_{E_{Dx}}.$$

Let  $w_x \neq 0$  be such that  $w_x \in \ker \omega_D(x) \cap W_{cx}$ . Then  $v_x + w_x$  also satisfies (5.11), which shows that the affine distribution described above has dimension greater than 0. Using this, the proof of the equivalence between uniqueness of solution and (A)(i) follows easily. The rest of the proof of the equivalence with (A) follows from the fact that  $\ker \omega_D(x) = E_{Dx}^{\omega_D} = E_{Dx}^\Omega \cap E_{Dx}$  for each  $x \in M_c$ , which can be proved directly using the definitions.

Note that the equivalence between (i), (ii), (iii) of (A) holds for *any* subspace  $W_{cx} \subseteq E_{Dx}$ .

Now we shall prove (B). First of all, if  $v_x \in W_{cx}$  satisfies (5.11) then, since  $W_{cx} \subseteq E_{Dx}$ , it clearly also satisfies

$$(\omega_D(x)|_{W_{cx}})(v_x, \cdot) = d\mathcal{E}|_{W_{cx}}. \quad (5.13)$$

We can conclude that if  $\omega_D(x)|_{W_{cx}}$ , which as we know coincides with  $\Omega(x)|_{W_{cx}}$ , is symplectic then equation (5.11) defines a vector field  $v_x$  on  $M_c$ , which is, in fact, given by equation (5.13). Then existence and uniqueness of solution  $x(t)$  of (5.11) for any initial condition  $x(0) \in M_c$  is guaranteed. Finally, the equivalence between symplecticity of  $\omega_D(x)|_{W_{cx}} \equiv \Omega(x)|_{W_{cx}}$  and (B)(i) and (B)(ii) is easy to prove using basic linear symplectic geometry. In fact, it is easy to prove that symplecticity of  $\omega_D(x)|_{W_{cx}}$  is equivalent to  $W_{cx}^{\omega_D} = \{0\}$ . Using Lemma 3.4 we can deduce that  $W_{cx}^{\omega_D} = W_{cx} \cap W_{cx}^\Omega$  from which we obtain  $W_{cx} \cap W_{cx}^\Omega = W_{cx} \cap W_{cx}^{\omega_D}$ . ■

With the method just described, one can deal with many examples of interest, such as nonholonomic systems and circuits, provided that one chooses the manifold  $M$  and the Dirac structure  $D$  properly. We show in the next section how a nonholonomic system given by a distribution  $\Delta \subseteq TQ$  on the configuration space  $Q$  can be described by a Dirac dynamical system on the Pontryagin bundle  $M = TQ \oplus T^*Q$

and how one can apply the constraint algorithm CAD to this example, although we will not perform a detailed calculation of the sequence of constraint submanifolds  $M_k$ . We also show how LC circuits can be treated with the same formalism as nonholonomic systems. The main point for doing this is, again, to choose the manifold  $M$  as being the Pontryagin bundle  $TQ \oplus T^*Q$ , where, this time,  $Q$  is the charge space, and a canonically constructed Dirac structure  $\bar{D}_\Delta$  on  $M$ , where, this time,  $\Delta$  represents Kirchhoff's Current Law. We also show how this approach using  $TQ \oplus T^*Q$  is related to the approach used in [Yoshimura and Marsden \[2006a,b\]](#).

## 6 Examples

In this section we deal with two examples, namely, nonholonomic systems, and LC circuits, showing that Dirac dynamical systems give a unified treatment for them. We will perform the detailed calculation of the constraint submanifolds  $M_k$  for the case of LC circuits only.

**Nonholonomic systems.** Recall that a nonholonomic system is given by a configuration space  $Q$ , a distribution  $\Delta \subseteq TQ$ , called the nonholonomic constraint, and a Lagrangian  $\mathcal{L}: TQ \rightarrow \mathbb{R}$ . Equations of motion are given by Lagrange-d'Alembert's principle.

Inspired by the Hamilton-Poincaré principle given in [Cendra, Marsden, Pekarsky, and Ratiu \[2003a\]](#), we can write a convenient equivalent form of the Lagrange-d'Alembert principle as

$$\delta \int_{t_0}^{t_1} (p\dot{q} - \mathcal{E}(q, v, p)) dt = 0,$$

where  $\mathcal{E}: TQ \oplus T^*Q \rightarrow \mathbb{R}$  is defined by  $\mathcal{E}(q, v, p) = pv - \mathcal{L}(q, v)$ , and with the restriction on variations  $\delta q \in \Delta$ ,  $\delta q(t_i) = 0$  for  $i = 0, 1$ , along with the kinematic restriction  $v \in \Delta$ . The resulting equations are

$$\dot{p} - \frac{\partial \mathcal{L}}{\partial q} \in \Delta^\circ \tag{6.1}$$

$$\dot{q} = v \tag{6.2}$$

$$p - \frac{\partial \mathcal{L}}{\partial v} = 0 \tag{6.3}$$

$$v \in \Delta. \tag{6.4}$$

We are going to show that equations (6.1)–(6.4) can be written in the form (1.1). For this purpose we must construct an appropriate Dirac structure associated to the nonholonomic constraint. Inspired by several results in [Yoshimura and Marsden \[2006a\]](#) and by the Hamilton-Poincaré's point of view we define a Dirac structure  $\bar{D}_\Delta \subseteq TM \oplus T^*M$  on  $M = TQ \oplus T^*Q$  associated to a given distribution  $\Delta \subseteq TQ$  on a manifold  $Q$  by the local expression

$$\bar{D}_\Delta(q, v, p) = \{(q, v, p, \dot{q}, \dot{v}, \dot{p}, \alpha, \gamma, \beta) \mid \dot{q} \in \Delta(q), \alpha + \dot{p} \in \Delta^\circ(q), \beta = \dot{q}, \gamma = 0\}.$$

**Note.** We shall accept both equivalent notations

$$(q, v, p, \dot{q}, \dot{v}, \dot{p}, \alpha, \gamma, \beta) \equiv (q, v, p, \dot{q}, \dot{v}, \dot{p}) \oplus (q, v, p, \alpha, \gamma, \beta),$$

for an element of  $TM \oplus T^*M$ .

By checking that  $\dim \bar{D}_\Delta(q, v, p) = 3 \dim Q$ , that is  $\dim \bar{D}_\Delta(q, v, p) = \dim M$ , and that

$$\langle\langle (q, v, p, \dot{q}_1, \dot{v}_1, \dot{p}_1) \oplus (q, v, p, \alpha_1, \gamma_1, \beta_1), (q, v, p, \dot{q}_2, \dot{v}_2, \dot{p}_2) \oplus (q, v, p, \alpha_2, \gamma_2, \beta_2) \rangle\rangle = 0$$

for all

$$\begin{aligned} (q, v, p, \dot{q}_1, \dot{v}_1, \dot{p}_1) \oplus (q, v, p, \alpha_1, \gamma_1, \beta_1) &\in \bar{D}_\Delta(q, v, p), \\ (q, v, p, \dot{q}_2, \dot{v}_2, \dot{p}_2) \oplus (q, v, p, \alpha_2, \gamma_2, \beta_2) &\in \bar{D}_\Delta(q, v, p), \end{aligned}$$

we can conclude using Lemma 4.1 that  $\bar{D}_\Delta$  is a Dirac structure on  $M$ . We should now prove that  $\bar{D}_\Delta$  is well defined globally, in other words, that it does not depend on the choice of a local chart. Let  $\bar{\tau}: TQ \oplus T^*Q \rightarrow Q$  and  $\bar{\pi}: TQ \oplus T^*Q \rightarrow T^*Q$  be the natural maps that in local coordinates are given by  $\bar{\tau}(q, v, p) = q$  and  $\bar{\pi}(q, v, p) = (q, p)$ . For a given distribution  $\Delta \subseteq TQ$  consider the distribution  $\bar{\Delta} = (T\bar{\tau})^{-1}(\Delta)$  and also the 2-form  $\bar{\omega} = \bar{\pi}^*\omega$ , on the manifold  $TQ \oplus T^*Q$ , where  $\omega$  is the canonical 2-form on  $T^*Q$ . We have the local expressions  $\bar{\Delta} = \{(q, v, p, \dot{q}, \dot{v}, \dot{p}) \mid \dot{q} \in \Delta\}$  and  $\bar{\omega}(q, v, p) = dq \wedge dp$ . Now we can apply Theorem 4.5 replacing  $M$  by  $TQ \oplus T^*Q$ ,  $E$  by  $\bar{\Delta}$  and  $\omega$  by  $\bar{\omega}$  and then we can easily check that the Dirac structure  $D_{\omega_E}$  coincides with  $\bar{D}_\Delta$ . In other words, by using Theorem 4.5 we have obtained a coordinate-independent description of  $\bar{D}_\Delta$  given in terms of  $\bar{\tau}$ ,  $\bar{\pi}$  and  $\bar{\omega}$ , which proves in particular that it is well defined globally.

It is straightforward to check that the condition

$$(x, \dot{x}) \oplus d\mathcal{E}(x) \in \bar{D}_\Delta, \quad (6.5)$$

where  $x = (q, v, p)$ , is equivalent to

$$\dot{p} - \frac{\partial \mathcal{L}}{\partial q} \in \Delta^\circ \quad (6.6)$$

$$\dot{q} = v \quad (6.7)$$

$$p = \frac{\partial \mathcal{L}}{\partial v} \quad (6.8)$$

$$\dot{q} \in \Delta, \quad (6.9)$$

which is clearly equivalent to equations (6.1)–(6.4).

Now that we have written the equations of motion as a Dirac dynamical system, we can proceed to apply the CAD. First, we can easily prove the following formulas using the definitions and Proposition 4.4:

$$E_{\bar{D}_\Delta} = \{(q, v, p, \dot{q}, \dot{v}, \dot{p}) \mid \dot{q} \in \Delta\}, \quad (6.10)$$

$$E_{\bar{D}_\Delta}^\flat = \{(q, v, p, \alpha, \gamma, \beta) \mid \gamma = 0, \beta \in \Delta\}, \quad (6.11)$$

$$E_{\bar{D}_\Delta}^{\bar{D}_\Delta} = \{(q, v, p, \dot{q}, \dot{v}, \dot{p}) \mid \dot{q} = 0, \dot{p} \in \Delta^\circ\}. \quad (6.12)$$

Then we have

$$\begin{aligned}
M_1 &= \left\{ (q, v, p) \mid \left\langle d\mathcal{E}(q, v, p), E_{D_\Delta}^{\bar{D}_\Delta}(q, v, p) \right\rangle = \{0\} \right\} \\
&= \left\{ (q, v, p) \mid d\mathcal{E}(q, v, p) \in E_{D_\Delta}^b(q, v, p) \right\} \\
&= \left\{ (q, v, p) \mid p - \frac{\partial L}{\partial v} = 0, v \in \Delta \right\}.
\end{aligned} \tag{6.13}$$

We could continue applying the algorithm as explained in general in section 5, and we would obtain specific formulas for  $M_2, M_3$ , etc.

**Relationship with implicit Lagrangian systems.** In Yoshimura and Marsden [2006a] the description of a nonholonomic system as an *implicit Lagrangian system* was introduced and equations of motion were shown to be a *partial vector field* written in terms of the *Dirac differential*. There is a close and simple relationship between this approach and the one of the present paper, which we shall explain next.

First, we should recall the notion of an implicit Lagrangian system. Let  $\mathcal{L}: TQ \rightarrow \mathbb{R}$  be a given Lagrangian; then  $d\mathcal{L}$  is a 1-form on  $TQ$ ,  $d\mathcal{L}: TQ \rightarrow T^*TQ$ , which is locally expressed by

$$d\mathcal{L} = \left( q, v, \frac{\partial \mathcal{L}}{\partial q}, \frac{\partial \mathcal{L}}{\partial v} \right).$$

Define a differential operator  $\mathfrak{D}$ , called the *Dirac differential* of  $\mathcal{L}$ , by

$$\mathfrak{D}\mathcal{L} = \gamma_Q \circ d\mathcal{L},$$

where  $\gamma_Q$  is the diffeomorphism defined by

$$\gamma_Q = \Omega^b \circ (\kappa_Q)^{-1}: T^*TQ \rightarrow T^*T^*Q.$$

Here  $\Omega$  is the canonical 2-form on  $T^*Q$  and  $\kappa_Q: TT^*Q \rightarrow T^*TQ$  is the canonical isomorphism which is given in a local chart by  $\kappa_Q(q, p, \delta q, \delta p) = (q, \delta q, \delta p, p)$ . We have the local expression

$$\mathfrak{D}\mathcal{L} = \left( q, \frac{\partial \mathcal{L}}{\partial v}, -\frac{\partial \mathcal{L}}{\partial q}, v \right).$$

We must now recall the definition and properties of the Dirac structure  $D_\Delta$  on the manifold  $T^*Q$ , where  $\Delta$  is a given distribution on  $Q$ , studied in propositions 5.1 and 5.2 of Yoshimura and Marsden [2006a]. Let  $\pi: T^*Q \rightarrow Q$  be the canonical projection. Then  $\Delta_{T^*Q}$  is defined by

$$\Delta_{T^*Q} = (T\pi)^{-1}(\Delta).$$

The Dirac structure  $D_\Delta$  is the one given by Theorem 4.5, with  $M = T^*Q$ ,  $\omega = \Omega$  the canonical symplectic form on  $T^*Q$  and  $E = \Delta_{T^*Q}$ , and it is described by

$$\begin{aligned}
D_\Delta(q, p) &= \{v \oplus \alpha \in T_{(q,p)}T^*Q \oplus T_{(q,p)}^*T^*Q \mid v \in \Delta_{T^*Q}(q, p) \\
&\text{and } \alpha(w) = \Omega(q, p)(v, w) \text{ for all } w \in \Delta_{T^*Q}(q, p)\}.
\end{aligned} \tag{6.14}$$

Now we will define the notion of an implicit Lagrangian system.



**Definition 6.1.** Let  $\mathcal{L}: TQ \rightarrow \mathbb{R}$  be a given Lagrangian (possibly degenerate) and let  $\Delta \subseteq TQ$  be a given regular constraint distribution on a configuration manifold  $Q$ . Denote by  $D_\Delta$  the induced Dirac structure on  $T^*Q$  that is given by the equation (6.14) and write  $\mathfrak{D}: TQ \rightarrow T^*T^*Q$  for the Dirac differential of  $\mathcal{L}$ . Let  $\mathbb{F}L(\Delta) \subseteq T^*Q$  be the image of  $\Delta$  under the Legendre transformation. An implicit Lagrangian system is a triple  $(\mathcal{L}, \Delta, X)$  where  $X$  represents a vector field defined at points of  $\mathbb{F}L(\Delta) \subseteq T^*Q$ , together with the condition

$$X \oplus \mathfrak{D}\mathcal{L} \in D_\Delta.$$

In other words, for each point  $(q, v) \in \Delta$  let  $(q, p) = \mathbb{F}L(q, v)$  and then, by definition,  $X(q, v, p) \in T_{(q,p)}(T^*Q)$  must satisfy

$$X(q, v, p) \oplus \mathfrak{D}\mathcal{L}(q, v) \in D_\Delta(q, p). \quad (6.15)$$

A solution curve of an implicit Lagrangian system  $(\mathcal{L}, \Delta, X)$  is a curve  $(q(t), v(t)) \in \Delta$ ,  $t_1 \leq t \leq t_2$ , which is an integral curve of  $X$  where  $(q(t), p(t)) = \mathbb{F}L(q(t), v(t))$ . The following proposition is essentially Proposition 6.3 of Yoshimura and Marsden [2006a].

**Proposition 6.2.** The condition  $X \oplus \mathfrak{D}\mathcal{L} \in D_\Delta$  defining an implicit Lagrangian system is given locally by the equalities

$$p = \frac{\partial \mathcal{L}}{\partial v} \quad (6.16)$$

$$\dot{q} \in \Delta \quad (6.17)$$

$$v = \dot{q} \quad (6.18)$$

$$\dot{p} - \frac{\partial \mathcal{L}}{\partial q} \in \Delta^\circ. \quad (6.19)$$

It is clear that equations (6.16)–(6.19) are equivalent to (6.1)–(6.4) and also to (6.6)–(6.9). This leads immediately to a precise link between the approach to nonholonomic systems given in Yoshimura and Marsden [2006a] and the one in the present paper, which is given in the next proposition whose proof in local coordinates is easy and will be omitted. Recall that  $M_1$  is the first constraint manifold given by (6.13).

**Proposition 6.3.** Let  $\mathcal{L}: TQ \rightarrow \mathbb{R}$  be a given Lagrangian,  $\Delta \subseteq TQ$  a given distribution and let  $\mathcal{E}: TQ \oplus T^*Q \rightarrow \mathbb{R}$  be given by  $\mathcal{E}(q, v, p) = pv - \mathcal{L}(q, v)$ . Then the following assertions hold.

- (i)  $\mathbb{F}L(\Delta) = \pi(M_1)$ .
- (ii) Let  $(q, v, p, \dot{q}, \dot{v}, \dot{p}) = \bar{X}(q, v, p)$  be a solution to (6.5); then in particular one has  $(q, v, p) \in M_1$ , and let  $X(q, v, p) = (q, p, \dot{q}, \dot{p})$ , that is,  $X(q, v, p) = T\pi\bar{X}(q, v, p)$ . Then  $X(q, v, p)$  is a solution to (6.15).
- (iii) A curve  $(q(t), v(t), p(t))$  is a solution curve of (6.5) if and only if  $(q(t), v(t))$  is a solution curve of the implicit Lagrangian system given by (6.15).

**LC circuits.** This is a case of a constant Dirac structure of the type explained in section 5. Our approach is closely related and equivalent to the one described in Yoshimura and Marsden [2006a]. There are many relevant references where the structure of Kirchhoff's laws has been studied from different points of view; some of them emphasize the geometry behind the equations, see Smale [1972]; Chua and McPherson [1974]; Maschke, van der Schaft, and Breedveld [1995]; Moreau and Aeyels [2004].

We first briefly recall the description of LC circuits given in Yoshimura and Marsden [2006a]. Let  $E$  be a vector space representing the *charge space*; then  $TE$  is the *current space* and  $V = T^*E$  is the *flux linkage space*. There is a constant distribution  $\Delta \subset TE$  (that is,  $\Delta$  is invariant under translations), which represents the *Kirchhoff's Current Law* (KCL). The subbundle  $\Delta^\circ \subseteq T^*E$  represents the *Kirchhoff's Voltage Law* (KVL). One has a Dirac structure on the cotangent bundle  $T^*E$ ,  $D_\Delta \subseteq TV \oplus T^*V$  given by

$$D_\Delta = \{(q, p, \dot{q}, \dot{p}) \oplus (q, p, \alpha_q, \alpha_p) \in TV \oplus T^*V \mid \dot{q} \in \Delta, \alpha_p = \dot{q}, \alpha_q + \dot{p} \in \Delta^\circ\},$$

which clearly does not depend on  $(q, p)$ , in other words it is a constant Dirac structure, so for each base point  $(q, p)$  we have

$$D_\Delta(q, p) = (q, p, \tilde{D}),$$

where

$$\tilde{D} = \{(\dot{q}, \dot{p}, \alpha, \beta) \in V \oplus V^* \mid \dot{q} \in \Delta, \beta = \dot{q}, \alpha + \dot{p} \in \Delta^\circ\}.$$

The dynamics of the system is given by a Lagrangian  $\mathcal{L}: TE \rightarrow \mathbb{R}$ . This Lagrangian is given by a quadratic form on  $E \times E$  representing the difference of the energies in the inductors and the capacitors, say,

$$\mathcal{L}(q, v) = \frac{1}{2} \sum_{i=1}^n L_i v_i^2 - \frac{1}{2} \sum_{i=1}^n \frac{1}{C_i} q_i^2,$$

where  $n$  is the number of branches of the circuit. Of course, some of the terms in the previous sum may be zero, corresponding to the absence of an inductor or a capacitor in the corresponding branch.

One approach would be to use the time evolution of the circuit in terms of the Dirac differential, given by equation (6.15). This shows that Dirac structures provide a unified treatment for nonholonomic systems and LC circuits. In other words the equation is

$$(\mathbb{F}\mathcal{L}(q, v), \dot{q}, \dot{p}) \oplus (q, p, \mathfrak{D}\mathcal{L}(q, v)) \in D_\Delta(q, p),$$

with  $(q, p) = \mathbb{F}\mathcal{L}(q, v)$  and  $(q, v) \in \Delta$ , where  $\mathfrak{D}\mathcal{L}$  represents the Dirac differential of  $\mathcal{L}$ .

However, we will follow the philosophy of this paper and choose to work on  $TE \oplus T^*E$  rather than working with the Dirac differential.

**Applying the CAD to LC circuits.** It should be now clear how we can apply the constraint algorithm CAD developed in the present paper to deal with LC circuits exactly as we did with the case of nonholonomic systems. More precisely, it should now be clear that an LC circuit can be described by the methods described in the first part of the paragraph *Nonholonomic Systems* (beginning of section 6), by taking  $Q = E$ ,  $M = TE \oplus T^*E$  and defining  $\bar{D}_\Delta$  and  $\mathcal{E}$  as indicated in that paragraph. As we have already said, in the case of circuits the Dirac structure  $\bar{D}_\Delta$  is constant and therefore integrable. However, we prefer not to work with the system restricted to a presymplectic leaf, as explained before, but to apply directly the algorithm for a general Dirac structure, as we will explain next.

Define the linear maps  $\varphi: E \rightarrow E^*$  and  $\psi: E \rightarrow E^*$  by

$$\begin{aligned}\varphi(v) &= \frac{\partial \mathcal{L}}{\partial v} = (L_1 v_1, \dots, L_n v_n), \\ \psi(q) &= \frac{\partial \mathcal{L}}{\partial q} = -(q_1/C_1, \dots, q_n/C_n).\end{aligned}$$

If there is no capacitor in a circuit branch, the corresponding component of  $\psi(q)$  will be zero (infinite capacitance), while zero capacitance is ruled out since it would represent an electrically open circuit branch. The physically relevant case corresponds to  $0 \leq L_i < \infty$ ,  $0 < C_i \leq \infty$ , for  $i = 1, \dots, n$ . However, the constraint algorithm applies to the general case where capacitances and inductances can be negative.

The evolution equations (6.6)–(6.9) for a general nonholonomic system become

$$\dot{p} - \psi(q) \in \Delta^\circ \quad (6.20)$$

$$\dot{q} = v \quad (6.21)$$

$$p = \varphi(v) \quad (6.22)$$

$$\dot{q} \in \Delta. \quad (6.23)$$

Now we will apply the CAD. First, note that

$$d\mathcal{E}(q, v, p) = \psi(q)dq + (p - \varphi(v))dv + vdp.$$

The first constraint submanifold is calculated as follows, taking into account the general expression (6.11),

$$\begin{aligned}M_1 &= \{(q, v, p) \in M \mid d\mathcal{E}(q, v, p) \in E_{\bar{D}_\Delta}^b(q, v, p)\} \\ &= \{(q, v, p) \mid p = \varphi(v), v \in \Delta\}.\end{aligned}$$

Now we simply continue applying the algorithm. Let

$$W_1 = TM_1 \cap E_{\bar{D}_\Delta} = \{(q, v, p, \dot{q}, \dot{v}, \dot{p}) \mid (q, v, p) \in M_1, \dot{q} \in \Delta, \dot{p} = \varphi(\dot{v}), \dot{v} \in \Delta\}.$$

Then  $\bar{D}_\Delta^b(W_1)$ , denoted simply as  $W_1^b$ , is

$$W_1^b = \{(q, v, p, \alpha, \gamma, \beta) \mid (q, v, p) \in M_1, \alpha \in \varphi(\Delta) + \Delta^\circ, \gamma = 0, \beta \in \Delta\},$$

and therefore

$$\begin{aligned} M_2 &= \{(q, v, p) \in M_1 \mid d\mathcal{E}(q, v, p) \in W_1^b(q, v, p)\} \\ &= \{(q, v, p) \mid q \in \psi^{-1}(\varphi(\Delta) + \Delta^\circ), p = \varphi(v), v \in \Delta\}. \end{aligned} \quad (6.24)$$

In the same way we can calculate  $M_3$ . In fact,

$$\begin{aligned} W_2 &= TM_2 \cap E_{\bar{D}_\Delta} = \{(q, v, p, \dot{q}, \dot{v}, \dot{p}) \mid (q, v, p) \in M_2, \dot{q} \in \Delta, \dot{q} \in \psi^{-1}(\varphi(\Delta) + \Delta^\circ), \\ &\quad \dot{p} = \varphi(\dot{v}), \dot{v} \in \Delta\}, \end{aligned}$$

$$W_2^b = \{(q, v, p, \alpha, \gamma, \beta) \mid (q, v, p) \in M_2, \alpha \in \varphi(\Delta) + \Delta^\circ, \gamma = 0, \beta \in \Delta \cap \psi^{-1}(\varphi(\Delta) + \Delta^\circ)\},$$

then

$$\begin{aligned} M_3 &= \{(q, v, p) \in M_2 \mid d\mathcal{E}(q, v, p) \in W_2^b(q, v, p)\} \\ &= \{(q, v, p) \mid q \in \psi^{-1}(\varphi(\Delta) + \Delta^\circ), p = \varphi(v), v \in \Delta \cap \psi^{-1}(\varphi(\Delta) + \Delta^\circ)\} \\ &= \{(q, v, p) \mid q \in \psi^{-1}(\varphi(\Delta) + \Delta^\circ), p = \varphi(v), v \in \Delta_1\}, \end{aligned}$$

where we have called

$$\Delta_1 = \Delta \cap \psi^{-1}(\varphi(\Delta) + \Delta^\circ) \subseteq \Delta.$$

Now we shall calculate  $M_4$ . We will show later that  $M_4 = M_3$  for all LC circuits with either a positive inductance or a non-infinite positive capacitance on every branch (Lemma 6.5 and Theorem 6.6). First we compute

$$\begin{aligned} W_3 &= TM_3 \cap E_{\bar{D}_\Delta} \\ &= \{(q, v, p, \dot{q}, \dot{v}, \dot{p}) \mid (q, v, p) \in M_3, \dot{q} \in \Delta, \dot{q} \in \psi^{-1}(\varphi(\Delta) + \Delta^\circ), \dot{p} = \varphi(\dot{v}), \dot{v} \in \Delta_1\} \\ &= \{(q, v, p, \dot{q}, \dot{v}, \dot{p}) \mid (q, v, p) \in M_3, \dot{q} \in \Delta_1, \dot{p} = \varphi(\dot{v}), \dot{v} \in \Delta_1\}. \end{aligned}$$

Thus, one finds that

$$W_3^b = \{(q, v, p, \alpha, \gamma, \beta) \mid (q, v, p) \in M_3, \alpha \in \varphi(\Delta_1) + \Delta^\circ, \gamma = 0, \beta \in \Delta_1\},$$

and hence

$$\begin{aligned} M_4 &= \{(q, v, p) \in M_3 \mid d\mathcal{E}(q, v, p) \in W_3^b(q, v, p)\} \\ &= \{(q, v, p) \in M_3 \mid q \in \psi^{-1}(\varphi(\Delta_1) + \Delta^\circ), p = \varphi(v), v \in \Delta_1\} \\ &= \{(q, v, p) \mid q \in \psi^{-1}(\varphi(\Delta_1) + \Delta^\circ), p = \varphi(v), v \in \Delta_1\}. \end{aligned}$$

Similarly we calculate  $M_5$  as follows.

$$\begin{aligned} W_4 &= TM_4 \cap E_{\bar{D}_\Delta} \\ &= \{(q, v, p, \dot{q}, \dot{v}, \dot{p}) \mid (q, v, p) \in M_4, \dot{q} \in \Delta, \dot{q} \in \psi^{-1}(\varphi(\Delta_1) + \Delta^\circ), \dot{p} = \varphi(\dot{v}), \dot{v} \in \Delta_1\} \\ &= \{(q, v, p, \dot{q}, \dot{v}, \dot{p}) \mid (q, v, p) \in M_4, \dot{q} \in \Delta \cap \psi^{-1}(\varphi(\Delta_1) + \Delta^\circ), \dot{p} = \varphi(\dot{v}), \dot{v} \in \Delta_1\} \end{aligned}$$

Let  $\Delta_2 = \Delta \cap \psi^{-1}(\varphi(\Delta_1) + \Delta^\circ)$ ; then  $\Delta_2 \subseteq \Delta_1$  and we can rewrite  $W_4$  as

$$W_4 = \{(q, v, p, \dot{q}, \dot{v}, \dot{p}) \mid (q, v, p) \in M_4, \dot{q} \in \Delta_2, \dot{p} = \varphi(\dot{v}), \dot{v} \in \Delta_1\}.$$

Then we have

$$W_4^b = \{(q, v, p, \alpha, \gamma, \beta) \mid (q, v, p) \in M_4, \alpha \in \varphi(\Delta_1) + \Delta^\circ, \gamma = 0, \beta \in \Delta_2\}$$

$$\begin{aligned} M_5 &= \{(q, v, p) \in M_4 \mid d\mathcal{E}(q, v, p) \in W_4^b(q, v, p)\} \\ &= \{(q, v, p) \in M_4 \mid q \in \psi^{-1}(\varphi(\Delta_1) + \Delta^\circ), p = \varphi(v), v \in \Delta_2\} \\ &= \{(q, v, p) \mid q \in \psi^{-1}(\varphi(\Delta_1) + \Delta^\circ), p = \varphi(v), v \in \Delta_2\}. \end{aligned}$$

From this, one sees how to recursively define  $M_k$ . For all  $k \geq 1$  define

$$\Delta_k = \Delta \cap \psi^{-1}(\varphi(\Delta_{k-1}) + \Delta^\circ),$$

where  $\Delta_0 = \Delta$  by definition. We have the following expressions for the constraint submanifolds  $M_k$ .

$$\begin{aligned} M_1 &= \{(q, v, p) \mid p = \varphi(v), v \in \Delta\} \\ M_{2k-1} &= \{(q, v, p) \mid q \in \psi^{-1}(\varphi(\Delta_{k-2}) + \Delta^\circ), p = \varphi(v), v \in \Delta_{k-1}\}, k \geq 2 \\ M_{2k} &= \{(q, v, p) \mid q \in \psi^{-1}(\varphi(\Delta_{k-1}) + \Delta^\circ), p = \varphi(v), v \in \Delta_{k-1}\}, k \geq 1. \end{aligned}$$

In order to solve the system we may apply the method to solve the general equation (5.12) explained before. It is clear that the parametrization  $x = x(z_1, \dots, z_r)$ , where  $r$  is the dimension of the final constraint  $M_c$ , can be chosen to be a linear map and the IDE (5.12) will be a linear system.

This system is of course still an IDE. One of the purposes of this paper is to write Hamilton's equations of motion, using an extension of the Dirac procedure extended to primary foliated constraint submanifolds, developed in section 7. For this purpose, we will need to consider  $M = TE \oplus T^*E$  as being the primary foliated constraint embedded naturally in the symplectic manifold  $T^*TE$ .

### Physical interpretation of the constraint equations for LC circuits.

The equations defining the constraint submanifolds have an interesting interpretation in circuit theory terms. From now on we will assume the physically meaningful situation where  $0 \leq L_i < \infty$ ,  $0 < C_i \leq \infty$ ,  $i = 1, \dots, n$ . We will show that in this case, the algorithm stops either at  $M_1$  or at  $M_3$ .

A circuit has an underlying directed graph, since each branch has a direction in which the current flow will be regarded as positive. A *loop* is a closed sequence of different adjacent branches. It also has a direction, which does not have to be compatible with the directions of the branches involved. The set of loop currents is a set of generators for the KCL subspace  $\Delta \subseteq E$ .

We shall call a circuit branch *inductive* (resp. *capacitive*) if there is an inductor (resp. capacitor) present on that branch. A loop will be called *inductive* (resp.

*capacitive*) if it has an inductor (resp. capacitor) on at least one of its branches. A branch or loop that is capacitive but not inductive will be called *purely capacitive*. A branch or loop that is neither inductive nor capacitive will be called *empty*, since no capacitors or inductors are present. Thus, a non-inductive branch or loop can be either empty or purely capacitive, and a purely capacitive loop must have at least one capacitor, no inductors, and possibly some empty branches.

The first one of the equations defining  $M_1$  is  $p = \varphi(v)$ , which is  $p_i = L_i v_i$ ,  $i = 1, \dots, n$ . The quantity  $L_i v_i$  is the voltage corresponding to an inductor through which the current is  $v_i$ . Then  $p_i$  can be interpreted as a time integral of the voltage on branch  $i$  due to the inductor in that branch (the *flux linkage* of the inductor). The second equation,  $v \in \Delta$ , is just KCL for the currents  $v$ .

The next constraint submanifold  $M_2$  incorporates the equation  $\psi(q) \in \varphi(\Delta) + \Delta^\circ$ , which represents the KVL equations for the purely capacitive loops, as explained in the next theorem.

**Theorem 6.4.** *The subspace  $\varphi(\Delta) + \Delta^\circ \subseteq E^*$  represents Kirchhoff's Voltage Law for non-inductive loops (KVLNI for short). That is, its elements are precisely those branch voltage assignments on the circuit that satisfy the subset of KVL equations corresponding to the non-inductive loops. The equation  $\psi(q) \in \varphi(\Delta) + \Delta^\circ$  represents the condition that on every purely capacitive loop, the branch voltages  $q_i/C_i$  of the corresponding capacitors satisfy KVL. We call these the KVLPC equations (KVL for purely capacitive loops), and they form a subset of the KVLNI equations.*

**Proof.** Let  $\{\bar{e}_1, \dots, \bar{e}_n\}$  be the basis of  $E$  where  $\bar{e}_i$  is associated to branch  $i$  of the circuit, and let  $\{\underline{e}^1, \dots, \underline{e}^n\}$  be its dual basis. Recall that the branch voltage assignments on the circuit are elements of  $E^*$ , and  $\Delta^\circ$  represents Kirchhoff's Voltage Law in the sense that each element of  $\Delta^\circ$  is a branch voltage assignment satisfying KVL. The KVL equations are linear equations on the branch voltages and are therefore represented by elements of  $E$ . In fact, they are precisely the elements of  $(\Delta^\circ)^\circ = \Delta$ .

Consider a maximal set of independent, non-inductive oriented loops on the circuit, labeled  $1, \dots, m$ . Each loop gives rise to a linear equation on the corresponding branch voltages according to KVL. These KVL equations are represented by  $\eta_1, \dots, \eta_m \in E$ , and are a subset of the full KVL equations for the circuit. That is, these linear equations hold on any branch voltage assignment that is compatible with KVL, and therefore  $\eta_1, \dots, \eta_m \in \Delta$ . Note that each  $\eta_j = \eta_j^i \bar{e}_i$  does not involve any inductive branches; that is, if  $L_i \neq 0$  then  $\eta_j^i = \eta_j(\underline{e}^i) = 0$ . Also, the KVLNI equations  $\eta_1, \dots, \eta_m$  can be seen as the KVL equations for the circuit that is obtained by removing the inductive branches from the given circuit. Let us call this *the non-inductive subcircuit*, whose charge space  $E_{\text{NI}}$  is a subspace of  $E$  in a natural way. Also,  $E_{\text{NI}}^* \subseteq E^*$  is generated by  $\{\underline{e}^i\}$  where  $i$  ranges over all non-inductive branches.

Let us denote  $\Delta_{\text{NI}} \subseteq E_{\text{NI}}$  the KCL distribution for the non-inductive subcircuit. Note that  $\Delta_{\text{NI}} \subseteq \Delta$ , since any KCL-compatible current assignment on the non-inductive subcircuit corresponds to a KCL-compatible current assignment on the original circuit for which no current flows through the inductive branches. Reciprocally, elements of  $\Delta$  that are zero on the inductive branches can be regarded

as elements of  $\Delta_{\text{NI}}$ . Note that  $\eta_1, \dots, \eta_m$  are loop currents for the non-inductive subcircuit, and they form a basis of  $\Delta_{\text{NI}}$ .

For any  $v = v^i \bar{e}_i \in E$ ,  $\varphi(v) = \sum L_i v^i \bar{e}^i$ . Then for each  $j = 1, \dots, m$ ,  $\eta_j(\varphi(v)) = \sum L_i v^i \eta_j(\bar{e}^i) = 0$ . This is true in particular for  $v \in \Delta$ , so  $\eta_j \in (\varphi(\Delta))^\circ$ . Therefore  $\eta_j \in (\varphi(\Delta))^\circ \cap \Delta = (\varphi(\Delta) + \Delta^\circ)^\circ$ . In other words, each  $\eta_1, \dots, \eta_m$ , seen as a linear equation on  $E^*$ , holds on  $\varphi(\Delta) + \Delta^\circ \subseteq E^*$ .

Let us see that  $\varphi(\Delta) + \Delta^\circ$  is precisely the vector subspace of  $E^*$  defined by the equations  $\eta_1, \dots, \eta_m$ . Let  $\rho = \rho^i \bar{e}_i \in (\varphi(\Delta) + \Delta^\circ)^\circ = (\varphi(\Delta))^\circ \cap \Delta$  be another equation that holds on  $\varphi(\Delta) + \Delta^\circ$ . Then  $0 = \rho(\varphi(\rho)) = \sum L_i (\rho^i)^2$ . Since all inductances are nonnegative, this implies that the components of  $\rho$  corresponding to the inductive branches must be zero. Also,  $\rho \in \Delta$ , therefore  $\rho \in \Delta_{\text{NI}}$ . This means that  $\rho$  is a linear combination of  $\eta_1, \dots, \eta_m$ .

The equation  $\psi(q) \in \varphi(\Delta) + \Delta^\circ$  means that  $\eta_j(\psi(q)) = 0$  for  $j = 1, \dots, m$ . Only those  $\eta_j$  corresponding to purely capacitive loops give a condition on  $q$ . The remaining  $\eta_k$  correspond to empty loops, so  $\eta_k(\psi(q)) = 0$  amounts to the equation  $0 = 0$ . ■

The third constraint submanifold  $M_3$  is obtained from  $M_2$  by incorporating the equation  $\psi(v) \in \varphi(\Delta) + \Delta^\circ$ , which means that the currents on the branches of the purely capacitive loops satisfy the same KVLPC equations as the charges. It is clear that it must hold when we consider the dynamics, which includes the equation  $\dot{q} = v$ .

Suppose that there exists at least one purely capacitive loop so that the algorithm does not stop at  $M_1$ . Then it is possible to show that  $M_3 \subsetneq M_2$ . In order to do that, consider a purely capacitive loop, which by definition must involve at least one capacitor, and recall that all the capacitances are positive. The corresponding KVLPC equation is represented by  $\eta \in E$ , whose components are  $\pm 1$  or  $0$ . As in the proof of the previous theorem,  $\eta \in (\Delta^\circ)^\circ = \Delta$ , so it can be interpreted as a nonzero loop current, that is, a current that is  $\pm 1$  on all the branches of that loop (depending on the relative orientation of the branches with respect to the loop) and zero on the remaining branches. Note that  $\psi(\eta) \in E^*$  does not satisfy KVLPC, since in particular  $\eta(\psi(\eta)) = \sum (\eta^i)^2 / C_i = \sum 1 / C_i > 0$ , where the last sum is over the branches in the chosen loop. Therefore  $\Delta \not\subseteq \psi^{-1}(\varphi(\Delta) + \Delta^\circ)$  and  $M_3 \subsetneq M_2$ .

Now we will prove the interesting fact that since inductances are greater than or equal to zero and capacitances are positive numbers or  $+\infty$  the algorithm always stops at  $M_3$  (or earlier). We shall start by proving the following lemma.

**Lemma 6.5.** (a)  $\delta_1 \in \psi^{-1}(\varphi(\Delta) + \Delta^\circ)$  is equivalent to the condition that there exists  $\delta_0 \in \Delta$  such that  $\psi(\delta_1) - \varphi(\delta_0) \in \Delta^\circ$ .

(b) The condition  $\Delta_1 = \Delta_2$  is equivalent to the following condition:

**Condition (\*)**

For any  $\delta \in \Delta$  there exist  $\delta_0, \delta_1 \in \Delta$  such that the following hold:

$$\begin{aligned}\varphi(\delta) - \varphi(\delta_1) &\in \Delta^\circ \\ \psi(\delta_1) - \varphi(\delta_0) &\in \Delta^\circ.\end{aligned}$$

Equivalently, we can say that for each  $\delta \in \Delta$  there exists  $\delta_1 \in \Delta_1$  such that  $\varphi(\delta) - \varphi(\delta_1) \in \Delta^\circ$ .

(c) Condition (\*) implies that if  $M_4 = M_5$  then  $M_3 = M_4$ .

**Proof.** The proof of (a) is immediate taking into account the definitions. To prove (b) take any given  $\delta_1 \in \Delta_1$ , then it satisfies the condition  $\psi(\delta_1) - \varphi(\delta_0) \in \Delta^\circ$ , where  $\delta_0 \in \Delta$ . By condition (\*) there exists  $\delta'_1 \in \Delta_1$  such that  $\varphi(\delta_0) - \varphi(\delta'_1) \in \Delta^\circ$ , and we can conclude that  $\psi(\delta_1) - \varphi(\delta'_1) \in \Delta^\circ$ , which shows that  $\delta_1 \in \Delta_2$ . The converse can be easily verified.

Now we shall prove (c). Since the inclusion  $M_4 \subseteq M_3$  is immediate by construction we will only prove the converse. Let  $(q, v, p) \in M_3$  then, in particular, there exists  $\delta \in \Delta$  such that  $\psi(q) - \varphi(\delta) \in \Delta^\circ$ . By condition (\*) there exists  $\delta_1 \in \Delta_1$  such that  $\varphi(\delta_1) - \varphi(\delta) \in \Delta^\circ$ , which implies that  $\psi(q) - \varphi(\delta_1) \in \Delta^\circ$ , that is,  $q \in \psi^{-1}(\varphi(\Delta_1) + \Delta^\circ)$ , from which we can deduce that  $M_3 \subseteq M_4$ . ■

The following theorem gives a sufficient condition under which condition (\*) in the previous lemma holds, which considers the standard physically meaningful case. Also, we assume for simplicity that there are no empty branches. We should mention that it can be proven that even with empty branches, the algorithm stops at  $M_1$  or  $M_3$ .

**Theorem 6.6.** Condition (\*) in Lemma 6.5 holds if in every branch there is either a positive inductance or a non-infinite positive capacitance.

**Proof.** Condition (\*) can be reformulated as follows:

For any  $\delta \in \Delta$  there exists  $\delta_0, \delta_1 \in \Delta$  such that the following equalities are satisfied

$$\langle \varphi(\delta_1), \delta'_1 \rangle = \langle \varphi(\delta), \delta'_1 \rangle \quad (6.25)$$

$$\langle (\psi + \varphi)(\delta_1), \delta'_0 \rangle - \langle \varphi(\delta_0), \delta'_0 \rangle = \langle \varphi(\delta), \delta'_0 \rangle, \quad (6.26)$$

for all  $\delta'_1, \delta'_0 \in \Delta$ . This formulation has the advantage that  $\nu := \psi + \varphi$  is nonsingular.

The basis  $\bar{e}_i$ ,  $i = 1, \dots, n$  of the charge space  $E$  defines naturally an Euclidean metric by the condition  $\langle \bar{e}_i, \bar{e}_j \rangle = \delta_{ij}$  and an identification  $E \equiv E^*$  by the condition  $\bar{e}_i = \underline{e}^i$ ,  $i = 1, \dots, n$ , in particular obtains  $\Delta^\circ \equiv \Delta^\perp$ . The linear map  $\nu$  is self-adjoint and positive definite while  $\varphi$  is self-adjoint and positive semi-definite. System (6.25)–(6.26) can be written in the form

$$\begin{aligned}P_\Delta \circ \varphi | \Delta(\delta_1) &= P_\Delta \circ \varphi | \Delta(\delta) \\ P_\Delta \circ \nu | \Delta(\delta_1) - P_\Delta \circ \varphi | \Delta(\delta_0) &= P_\Delta \circ \varphi | \Delta(\delta_0).\end{aligned}$$

where  $P_\Delta: E \rightarrow \Delta$  is the orthogonal projection on  $\Delta$ .



One can choose an orthonormal basis of  $\Delta$  and then the system of equations (6.25)–(6.26) can be written in matrix form. Moreover we have an orthogonal decomposition  $\Delta = \Delta^{(1)} + \Delta^{(2)}$ , where  $\Delta_1$  and  $\Delta_2$  are the image and the kernel of  $P_\Delta \circ \varphi|_\Delta$ . Each vector  $\delta \in \Delta$  is decomposed as  $\delta = \delta^{(1)} + \delta^{(2)}$ . We can choose the basis in such a way that the block decomposition of the matrix representing  $P_\Delta \circ \varphi|_\Delta$  has the form

$$\begin{bmatrix} \varphi^{(1,1)} & \varphi^{(1,2)} \\ \varphi^{(2,1)} & \varphi^{(2,2)} \end{bmatrix},$$

where  $\varphi^{(1,2)} = \varphi^{(2,1)} = \varphi^{(2,2)} = 0$  and  $\varphi^{(1,1)}$  is a diagonal matrix with positive eigenvalues. The map  $P_\Delta \circ \nu|_\Delta$  also has a block decomposition, and since it is self-adjoint and positive definite, the blocks  $\nu^{(1,1)}$  and  $\nu^{(2,2)}$  are symmetric and positive definite matrices. We obtain the following system of equations, which is equivalent to (6.25)–(6.26),

$$\begin{bmatrix} \varphi^{(1,1)} & \varphi^{(1,2)} \\ \varphi^{(2,1)} & \varphi^{(2,2)} \end{bmatrix} \begin{bmatrix} \delta_1^{(1)} \\ \delta_1^{(2)} \end{bmatrix} = \begin{bmatrix} \varphi^{(1,1)} & \varphi^{(1,2)} \\ \varphi^{(2,1)} & \varphi^{(2,2)} \end{bmatrix} \begin{bmatrix} \delta^{(1)} \\ \delta^{(2)} \end{bmatrix} \quad (6.27)$$

$$\begin{bmatrix} \nu^{(1,1)} & \nu^{(1,2)} \\ \nu^{(2,1)} & \nu^{(2,2)} \end{bmatrix} \begin{bmatrix} \delta_1^{(1)} \\ \delta_1^{(2)} \end{bmatrix} - \begin{bmatrix} \varphi^{(1,1)} & \varphi^{(1,2)} \\ \varphi^{(2,1)} & \varphi^{(2,2)} \end{bmatrix} \begin{bmatrix} \delta_0^{(1)} \\ \delta_0^{(2)} \end{bmatrix} = \begin{bmatrix} \varphi^{(1,1)} & \varphi^{(1,2)} \\ \varphi^{(2,1)} & \varphi^{(2,2)} \end{bmatrix} \begin{bmatrix} \delta^{(1)} \\ \delta^{(2)} \end{bmatrix} \quad (6.28)$$

For given  $\delta$ , equation (6.27) fixes  $\delta_1^{(1)}$  and imposes no condition on  $\delta_1^{(2)}$ . Using the fact that  $\varphi^{(1,1)}$  and  $\nu^{(2,2)}$  are invertible we can find  $\delta_0^{(1)}$  and  $\delta_1^{(2)}$  to satisfy (6.28). ■

As a conclusion, if (and only if) there are no purely capacitive loops in the circuit, then  $M_2 = M_1$  and the final constraint submanifold is  $M_1$ , defined by the conditions that  $p_i = L_i v_i$  (where some  $L_i$  might be zero) and  $v$  satisfies KCL. Otherwise, the final constraint submanifold is  $M_3$ , defined by the conditions that  $p_i = L_i v_i$ ,  $v$  satisfies KCL, and  $q_i/C_i$  and  $v_i/C_i$  satisfy KVLPC. Recall that  $q_i/C_i$  is the voltage of the capacitor in branch  $i$ , and the absence of a capacitor on branch  $i$  means  $C_i = \infty$ . The rest of the KVL equations, which involve the branch voltages  $\dot{p}_i = L_i \dot{v}_i$  on the inductors, will appear when considering the dynamics. In this sense, the KVLPC equations can be regarded as “static” KVL equations.

*Remark 6.7.* The hypothesis that no capacitances or inductances are negative, besides being physically meaningful, is crucial to ensure that the algorithm stops at  $M_3$  (or  $M_1$ ). For example, for a circuit with one inductor  $L$ , one capacitor  $C$  and a negative capacitor  $-C$ , all of them in parallel, it stops at  $M_5$ .

**Geometry of the final constraint submanifold.** We know that the solution curves will have tangent vectors in  $W_c = TM_c \cap E_{\bar{D}_\Delta}$ , where  $c = 1$  or  $c = 3$  depending on the case.

**Theorem 6.8.** *Assume the physically meaningful situation where  $0 \leq L_i < \infty$ ,  $0 < C_i \leq \infty$ ,  $i = 1, \dots, n$ . The distribution  $W_c \subseteq TM_c$  is constant and therefore*

integrable, and its integral leaves are preserved by the flow. Denote the leaf through 0 by  $\widetilde{W}_c$ , so these leaves can be written as  $x + \widetilde{W}_c$ , where  $x = (q, v, p)$ . Each integral leaf  $x + \widetilde{W}_c$  is a symplectic manifold with the pullback of the presymplectic form  $\omega = dq \wedge dp$ , if and only if there are no empty loops.

**Proof.** As mentioned before, if there are no purely capacitive loops, then  $c = 1$ , otherwise  $c = 3$ . For the first case, using (6.10), we have

$$\begin{aligned} W_1 &= \{(q, v, p, \dot{q}, \dot{v}, \dot{p}) \in TE \oplus T^*E \mid (q, v, p, \dot{q}, \dot{v}, \dot{p}) \in TM_1, \dot{q} \in \Delta\} \\ &= \{(q, v, p, \dot{q}, \dot{v}, \dot{p}) \in TE \oplus T^*E \mid p = \varphi(v), v \in \Delta, \dot{p} = \varphi(\dot{v}), \dot{v} \in \Delta, \dot{q} \in \Delta\} \end{aligned}$$

Then  $W_1 \subseteq TM_1$  is a constant and therefore integrable distribution whose rank is  $2 \dim \Delta$ . Write  $\omega_{x+\widetilde{W}_1}$  for the pullback of the presymplectic form  $\omega = dq \wedge dp$  to each integral leaf. Let us compute  $\ker \omega_{x+\widetilde{W}_1}$ .

Consider  $(q, v, p, \dot{q}, \dot{v}, \varphi(\dot{v})), (q, v, p, \dot{q}', \dot{v}', \varphi(\dot{v}')) \in W_1$ , where  $\dot{q}, \dot{v}, \dot{q}', \dot{v}' \in \Delta$  are arbitrary. Assume  $(q, v, p, \dot{q}, \dot{v}, \varphi(\dot{v})) \in \ker \omega_{x+\widetilde{W}_1}$ , that is,

$$0 = \omega((q, v, p, \dot{q}, \dot{v}, \varphi(\dot{v})), (q, v, p, \dot{q}', \dot{v}', \varphi(\dot{v}'))) = \dot{q}(\varphi(\dot{v}')) - \dot{q}'(\varphi(\dot{v}))$$

for any  $\dot{q}', \dot{v}' \in \Delta$ . This means that  $\dot{q}, \dot{v} \in (\varphi(\Delta))^\circ$ , but since they also belong to  $\Delta$  then  $\dot{q}, \dot{v} \in (\varphi(\Delta))^\circ \cap \Delta = (\varphi(\Delta) + \Delta^\circ)^\circ$ . That is,  $\ker \omega_{x+\widetilde{W}_1}$  is defined by  $\dot{q}, \dot{v} \in (\varphi(\Delta) + \Delta^\circ)^\circ, \dot{p} = \varphi(\dot{v})$ . The subspace  $(\varphi(\Delta) + \Delta^\circ)^\circ$  is generated by the loop currents on the non-inductive loops (Theorem 6.4). Since there are no purely capacitive loops, then it is generated by the loop currents on the empty loops. As a conclusion,  $x + \widetilde{W}_1$  is symplectic if and only if there are no empty loops on the circuit.

If there is at least one purely capacitive loop, then the final constraint submanifold is  $M_3$ . Then

$$\begin{aligned} W_3 &= \{(q, v, p, \dot{q}, \dot{v}, \dot{p}) \in TE \oplus T^*E \mid (q, v, p, \dot{q}, \dot{v}, \dot{p}) \in TM_3, \dot{q} \in \Delta\} \\ &= \{(q, v, p, \dot{q}, \dot{v}, \dot{p}) \in TE \oplus T^*E \mid p = \varphi(v), v \in \Delta, q \in \psi^{-1}(\varphi(\Delta) + \Delta^\circ), \\ &\quad v \in \psi^{-1}(\varphi(\Delta) + \Delta^\circ), \dot{p} = \varphi(\dot{v}), \dot{v} \in \Delta, \dot{q} \in \psi^{-1}(\varphi(\Delta) + \Delta^\circ), \\ &\quad \dot{v} \in \psi^{-1}(\varphi(\Delta) + \Delta^\circ), \dot{q} \in \Delta\}. \end{aligned}$$

Again,  $W_3 \subseteq TM_3$  is a constant and therefore integrable distribution whose rank is  $2 \dim(\Delta \cap \psi^{-1}(\varphi(\Delta) + \Delta^\circ)) = 2 \dim \Delta_1$ . Write  $\omega_{x+\widetilde{W}_3}$  for the pullback of the presymplectic form  $\omega = dq \wedge dp$  to each integral leaf, and let us compute  $\ker \omega_{x+\widetilde{W}_3}$ .

Consider  $(q, v, p, \dot{q}, \dot{v}, \varphi(\dot{v})), (q, v, p, \dot{q}', \dot{v}', \varphi(\dot{v}')) \in W_3$ , where  $\dot{q}, \dot{v}, \dot{q}', \dot{v}' \in \Delta_1$  are arbitrary. Assume  $(q, v, p, \dot{q}, \dot{v}, \varphi(\dot{v})) \in \ker \omega_{x+\widetilde{W}_3}$ , that is,

$$0 = \omega((q, v, p, \dot{q}, \dot{v}, \varphi(\dot{v})), (q, v, p, \dot{q}', \dot{v}', \varphi(\dot{v}'))) = \dot{q}(\varphi(\dot{v}')) - \dot{q}'(\varphi(\dot{v}))$$

for any  $\dot{q}', \dot{v}' \in \Delta_1$ . This means that  $\dot{q}, \dot{v} \in (\varphi(\Delta_1))^\circ$ . If there is an empty loop, represented by  $\eta \neq 0$ , then  $\eta \in \Delta, \psi(\eta) = 0$  and  $\eta \in \Delta_1$ . Also,  $\varphi(\eta) = 0$ , so  $(q, v, p, \eta, 0)$  is a nonzero element of  $\ker \omega_{x+\widetilde{W}_3}(q, v, p)$  and  $x + \widetilde{W}_3$  is not symplectic.

If there are no empty loops, we apply the following argument to  $\dot{q}$ , and the same conclusion holds for  $\dot{v}$ . Since  $\dot{q} \in (\varphi(\Delta_1))^\circ \cap \Delta_1$ , then  $\dot{q}(\varphi(\dot{q})) = 0$ , which means that the components of  $\dot{q}$  corresponding to the inductive branches are zero, as we reasoned in the proof of Theorem 6.4. Also,  $\dot{q} \in \Delta_1 \subset \Delta$ , so  $\dot{q} \in \Delta_{\text{NI}}$ , that is, it can be regarded as a branch current assignment on the non-inductive subcircuit, satisfying KCL. In addition,  $\dot{q}_i/C_i$  satisfy KVLNI, which means that  $\dot{q}_i/C_i$  satisfy KVL for the non-inductive subcircuit. Now we will apply the fact that the dynamics of a circuit with no inductors and no empty loops consists only of equilibrium points. Indeed, KVL for such a circuit is a homogeneous linear system on the charges  $q_i$  of the capacitors, so it has a solution  $\bar{q} = (\bar{q}_1, \dots, \bar{q}_n)$ . Setting  $q(t) = \bar{q}$  and  $v(t) = \dot{q}(t) = 0$ , we have that  $v$  satisfies KCL trivially. If there was a nonzero loop current  $\eta$ , then it would produce a change  $\dot{q} = \eta$  in the charges of the capacitor on that loop. Then they would no longer satisfy KVL for that loop, since  $d/dt(\eta(\psi(q))) = \eta(\psi(q)) \neq 0$ . Then the currents must therefore be zero, that is,  $\dot{q} = 0$ . Of course, this is not true if there are empty loops, which can have arbitrary loop currents. By the same reasoning,  $\dot{v} = 0$ , so  $\ker \omega_{x+\widetilde{W}_3} = 0$  and  $x + \widetilde{W}_3$  is symplectic. ■

Let us now write equations of motion in Hamiltonian form, by working on a symplectic leaf. This means that the initial conditions must belong to that particular leaf. This is related to the comments on equations (5.1) and (5.9) made before. An equation of motion in Poisson form that is valid for all initial conditions in  $M_c$  will be given in Theorem 7.4 by the equations (7.10) and (7.11).

First, let us represent the KCL equations by  $\kappa_1, \dots, \kappa_a \in E^*$ , in the sense that  $\kappa_i(v) = 0$  represents the KCL equation corresponding to node  $i$ . The number  $a$  is the number of independent nodes (usually the number of nodes minus one). As before, denote the KVLPC equations by  $\eta_1, \dots, \eta_b \in E$ , where  $b$  is the number of independent purely capacitive loops. Seeing each  $\kappa_i$  as a row vector and each  $\eta_j$  as a column vector, write

$$K = \begin{bmatrix} \kappa_1 \\ \vdots \\ \kappa_a \end{bmatrix}, \quad \Gamma = \begin{bmatrix} \eta_1 & \dots & \eta_b \end{bmatrix}.$$

The space  $TE \oplus T^*E$  with the presymplectic form  $\omega$  is embedded in  $T^*TE$  with the canonical symplectic form (see Appendix A), where the variables are  $(q, v, p, \nu)$ . The integral leaf  $x_0 + \widetilde{W}_c$ , where  $c = 1$  or  $c = 3$  as before, and  $x_0 = (q_0, v_0, p_0)$ , is defined as a subspace of  $T^*TE$  regularly by the equations

$$\begin{aligned} \kappa_i(q) - \kappa_i(q_0) &= 0, & i &= 1, \dots, a \\ \eta_i(\psi(q)) &= 0, & i &= 1, \dots, b \\ p_i - L_i v_i &= 0, & i &= 1, \dots, n \\ \kappa_i(v) &= 0, & i &= 1, \dots, a \\ \eta_i(\psi(v)) &= 0, & i &= 1, \dots, b \\ \nu_i &= 0, & i &= 1, \dots, n \end{aligned}$$

and the matrix of Poisson brackets of these constraints is, in block form,

$$\Sigma = \begin{bmatrix} 0 & 0 & K & 0 & 0 & 0 \\ 0 & 0 & \Gamma^T \psi & 0 & 0 & 0 \\ -K^T & -\psi \Gamma & 0 & 0 & 0 & -\varphi \\ 0 & 0 & 0 & 0 & 0 & K \\ 0 & 0 & 0 & 0 & 0 & \Gamma^T \psi \\ 0 & 0 & \varphi & -K^T & -\psi \Gamma & 0 \end{bmatrix},$$

where  $\varphi$  and  $\psi$  stand for the diagonal matrices with  $L_i$  and  $1/C_i$  along their diagonals, respectively.

**A concrete example.** We shall illustrate our method with the simple LC circuit studied in [Yoshimura and Marsden \[2006a\]](#) which is shown in Figure 6.1. It is a 4-port LC circuit where the configuration space is  $E = \mathbb{R}^4$ .

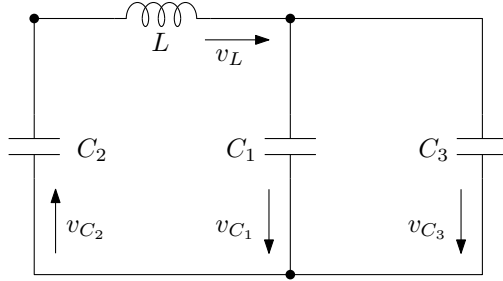


Figure 6.1: The constraint algorithm is applied to this 4-port LC circuit.

We shall use the notation  $q = (q_L, q_{C_1}, q_{C_2}, q_{C_3}) \in E$ ,  $v = (v_L, v_{C_1}, v_{C_2}, v_{C_3}) \in T_q E$ , and  $p = (p_L, p_{C_1}, p_{C_2}, p_{C_3}) \in T_q^* E$ . The Lagrangian of the LC circuit is  $\mathcal{L}: TE \rightarrow \mathbb{R}$ ,

$$\mathcal{L}(q, v) = \frac{1}{2} L (v_L)^2 - \frac{1}{2} \frac{(q_{C_1})^2}{C_1} - \frac{1}{2} \frac{(q_{C_2})^2}{C_2} - \frac{1}{2} \frac{(q_{C_3})^2}{C_3}.$$

We will assume the physically meaningful case where  $0 < L$  and  $0 < C_i < \infty$ ,  $i = 1, 2, 3$ . Then, by Theorem 6.6, the CAD algorithm will stop at most at  $M_3$ . Also, Theorem 6.4 gives a direct description of the specific equations that define  $M_1$ ,  $M_2$  and  $M_3$ . This circuit has a purely capacitive loop, so  $M_1 \supsetneq M_2 \supsetneq M_3$ . However, the algorithm could also be applied for the case where negative values for  $L$  or  $C_i$  are allowed. For this particular circuit, the algorithm stops at  $M_3$  regardless of the signs of the inductance and capacitances.

The KCL constraints  $\Delta \subseteq TE$  for the current  $v$  are

$$-v_L + v_{C_2} = 0 \tag{6.29}$$

$$-v_{C_1} + v_{C_2} - v_{C_3} = 0 \tag{6.30}$$

Therefore the constraint KCL space is defined, for each  $q \in E$ , by

$$\Delta(q) = \{v \in T_q E \mid \langle p^a, v \rangle = 0, a = 1, 2\},$$

where

$$(p_k^a) = \begin{pmatrix} -1 & 0 & 1 & 0 \\ 0 & -1 & 1 & -1 \end{pmatrix}.$$

On the other hand, the annihilator  $\Delta^\circ$  of  $\Delta$  is the constraint KVL space, defined, for each  $q \in E$ , by

$$\Delta^\circ(q) = \{p \in T_q^* E \mid \langle p, v \rangle = 0, \text{ for all } v \in \Delta(q)\}.$$

Note that  $\{p^a, a = 1, 2\}$  is a basis of  $\Delta^\circ$ . Taking into account that the Dirac structure  $\bar{D}_\Delta \subseteq TM \oplus T^*M$  on  $M = TE \oplus T^*E$  associated to the space  $\Delta(q)$  is, for each  $(q, v, p) \in M$ , given by

$$\begin{aligned} \bar{D}_\Delta(q, v, p) \\ = \{(q, v, p, \dot{q}, \dot{v}, \dot{p}, \alpha, \gamma, \beta) \in TM \oplus T^*M \mid \dot{q} \in \Delta(q), \alpha + \dot{p} \in \Delta^\circ(q), \beta = \dot{q}, \gamma = 0\}, \end{aligned}$$

and the corresponding energy  $\mathcal{E}: M \rightarrow \mathbb{R}$ , is given by  $\mathcal{E}(q, v, p) = pv - \mathcal{L}(q, v)$ , we can easily verify that the Dirac dynamical system

$$(q, v, p, \dot{q}, \dot{v}, \dot{p}) \oplus d\mathcal{E}(q, v, p) \in \bar{D}_\Delta(q, v, p)$$

is equivalent to the IDE system that is given in coordinates by (6.20)–(6.23), which gives

$$\dot{q}_L = v_L, \quad \dot{q}_{C_1} = v_{C_1}, \quad \dot{q}_{C_2} = v_{C_2}, \quad \dot{q}_{C_3} = v_{C_3} \quad (6.31)$$

$$\dot{p}_L + \frac{q_{C_1}}{C_1} + \dot{p}_{C_1} + \frac{q_{C_2}}{C_2} + \dot{p}_{C_2} = 0 \quad (6.32)$$

$$\dot{p}_L + \frac{q_{C_2}}{C_2} + \dot{p}_{C_2} + \frac{q_{C_3}}{C_3} + \dot{p}_{C_3} = 0 \quad (6.33)$$

$$p_L = Lv_L, \quad p_{C_1} = p_{C_2} = p_{C_3} = 0 \quad (6.34)$$

$$v_L = v_{C_2}, \quad v_{C_1} = v_{C_2} - v_{C_3}. \quad (6.35)$$

We now apply the constraint algorithm CAD for Dirac dynamical systems. We calculate the expressions of

$$\varphi(v) = \frac{\partial \mathcal{L}}{\partial v} \quad \text{and} \quad \psi(q) = \frac{\partial \mathcal{L}}{\partial q}$$

to get

$$\begin{aligned} \varphi(v_L, v_{C_1}, v_{C_2}, v_{C_3}) &= (Lv_L, 0, 0, 0) \\ \psi(q_L, q_{C_1}, q_{C_2}, q_{C_3}) &= (0, -\frac{q_{C_1}}{C_1}, -\frac{q_{C_2}}{C_2}, -\frac{q_{C_3}}{C_3}). \end{aligned}$$

From now on we will consider the constant distribution  $\Delta$  as a subspace of  $E$ . As we have seen before  $M_1 = \{(q, v, p) \mid v \in \Delta, p - \varphi(v) = 0\}$ , so using (6.29)–(6.30) and the expression for  $\varphi$  we get

$$M_1 = \{(q_L, q_{C_1}, q_{C_2}, q_{C_3}, v_L, v_{C_1}, v_{C_2}, v_{C_3}, p_L, p_{C_1}, p_{C_2}, p_{C_3}) \mid \\ p_L = Lv_L, p_{C_1} = p_{C_2} = p_{C_3} = 0, v_L = v_{C_2}, v_{C_1} = v_{C_2} - v_{C_3}\}.$$

We could calculate  $M_2$  using the expression (6.24). However, according to Theorem 6.4,  $M_2$  adds one more constraint, corresponding to the KVL for the purely capacitive loop that involves capacitors  $C_1$  and  $C_3$ . Then

$$M_2 = \left\{ (q_L, q_{C_1}, q_{C_2}, q_{C_3}, v_L, v_{C_1}, v_{C_2}, v_{C_3}, p_L, p_{C_1}, p_{C_2}, p_{C_3}) \mid p_L = Lv_L, \right. \\ \left. p_{C_1} = p_{C_2} = p_{C_3} = 0, v_L = v_{C_2}, v_{C_1} = v_{C_2} - v_{C_3}, \frac{q_{C_1}}{C_1} = \frac{q_{C_3}}{C_3} \right\}.$$

As explained right after the proof of Theorem 6.4, the final constraint submanifold  $M_3$  adds the constraint  $v_{C_1}/C_1 = v_{C_3}/C_3$ :

$$M_3 = \left\{ (q_L, q_{C_1}, q_{C_2}, q_{C_3}, v_L, v_{C_1}, v_{C_2}, v_{C_3}, p_L, p_{C_1}, p_{C_2}, p_{C_3}) \mid \right. \\ p_L = Lv_L, p_{C_1} = p_{C_2} = p_{C_3} = 0, v_L = v_{C_2}, v_{C_1} = v_{C_2} - v_{C_3}, \\ \left. \frac{q_{C_1}}{C_1} = \frac{q_{C_3}}{C_3}, \frac{v_{C_1}}{C_1} = \frac{v_{C_3}}{C_3} \right\}. \quad (6.36)$$

To solve the system we can simply parametrize  $M_3$ , which has dimension 4, for instance by taking some appropriate 4 of the 12 variables

$$(q_L, q_{C_1}, q_{C_2}, q_{C_3}, v_L, v_{C_1}, v_{C_2}, v_{C_3}, p_L, p_{C_1}, p_{C_2}, p_{C_3})$$

as being independent parameters and then replace in equations (6.31)–(6.33) to obtain an ODE equivalent to equations of motion. For instance, if we choose  $q_L, q_{C_1}, q_{C_2}, p_L$ , as independent variables we get the ODE

$$\begin{aligned} \dot{q}_L &= \frac{p_L}{L} \\ \dot{q}_{C_1} &= \frac{C_1}{L(C_1 + C_3)} p_L \\ \dot{q}_{C_2} &= \frac{p_L}{L} \\ \dot{p}_L &= -\frac{q_{C_1}}{C_1} - \frac{q_{C_2}}{C_2}. \end{aligned}$$

We know that solutions should be tangent to  $W_3 = TM_3 \cap E_{\tilde{D}_\Delta}$  where, according to (6.10),  $E_{\tilde{D}_\Delta}$  is defined by the conditions  $\dot{q} \in \Delta$ , that is,

$$\dot{q}_L = \dot{q}_{C_2}, \dot{q}_{C_1} = \dot{q}_{C_2} - \dot{q}_{C_3}. \quad (6.37)$$

Therefore  $W_3$  is defined by (6.37) together with the equations defining  $TM_3$  which are obtained by differentiating with respect to time the equations (6.36) defining  $M_3$ , that is,

$$\dot{p}_L = L\dot{v}_L, \dot{p}_{C_1} = \dot{p}_{C_2} = \dot{p}_{C_3} = 0, \dot{v}_L = \dot{v}_{C_2}, \dot{v}_{C_1} = \dot{v}_{C_2} - \dot{v}_{C_3}, \frac{\dot{q}_{C_1}}{C_1} = \frac{\dot{q}_{C_3}}{C_3}, \frac{\dot{v}_{C_1}}{C_1} = \frac{\dot{v}_{C_3}}{C_3}.$$

Then  $W_3$  is a constant and therefore integrable distribution of rank 2. Its integral leaves are preserved by the flow. One can check that the pullback of the presymplectic form  $\omega = dq \wedge dp$  to each leaf is nondegenerate, so they are symplectic manifolds. Denote  $\widetilde{W}_3 = W_3(0)$ , so these leaves can be written as  $x + \widetilde{W}_3$ .

For a given  $x_0 \in M_3$ , say

$$x_0 = (q_{L0}, q_{C_10}, q_{C_20}, q_{C_30}, v_{L0}, v_{C_10}, v_{C_20}, v_{C_30}, p_{L0}, p_{C_10}, p_{C_20}, p_{C_30})$$

we have that elements of the symplectic leaf  $x_0 + \widetilde{W}_3$  are characterized by the conditions defining  $M_3$  plus the condition obtained by integrating with respect to time the conditions (6.29) and (6.30), applied to  $(q_L - q_{L0}, q_{C_1} - q_{C_10}, q_{C_2} - q_{C_20}, q_{C_3} - q_{C_30})$ , that is

$$\begin{aligned} -(q_L - q_{L0}) + (q_{C_2} - q_{C_20}) &= 0 \\ -(q_{C_1} - q_{C_10}) + (q_{C_2} - q_{C_20}) - (q_{C_3} - q_{C_30}) &= 0. \end{aligned}$$

For simplicity, we will assume from now on that the conditions

$$\begin{aligned} q_{L0} - q_{C_20} &= 0 \\ q_{C_10} - q_{C_20} + q_{C_30} &= 0, \end{aligned}$$

are satisfied. While this is not the most general case, it is enough to illustrate the procedure.

We can conclude that  $x_0 + \widetilde{W}_3$  has dimension 2 and in fact the projection  $\bar{\pi}(x_0 + \widetilde{W}_3) \subseteq T^*E$  is simply the subspace of  $T^*E$  defined by the conditions  $q_{C_1} = q_{C_2} = q_{C_3} = 0$  and  $p_{C_1} = p_{C_2} = p_{C_3} = 0$ . Therefore one can use the variables  $(q_L, p_L)$  to parametrize  $\bar{\pi}(x_0 + \widetilde{W}_3)$ , and, in fact, we obtain

$$v_L = \frac{p_L}{L} \qquad q_{C_1} = \frac{C_1}{C_1 + C_3} q_L \qquad (6.38)$$

$$v_{C_1} = \frac{C_1}{C_1 + C_3} \frac{1}{L} p_L \qquad q_{C_2} = q_L \qquad (6.39)$$

$$v_{C_2} = \frac{p_L}{L} \qquad q_{C_3} = \frac{C_3}{C_1 + C_3} q_L \qquad (6.40)$$

$$v_{C_3} = \frac{C_3}{C_1 + C_3} \frac{1}{L} p_L \qquad p_{C_1} = p_{C_2} = p_{C_3} = 0 \qquad (6.41)$$

$$\nu_L = \nu_{C_1} = \nu_{C_2} = \nu_{C_3} = 0. \qquad (6.42)$$

Let  $x_0 + \widetilde{W}_3$  be defined by  $\epsilon_j = 0$ ,  $j = 1, \dots, 14$ , where we have chosen

$$\begin{aligned} \epsilon_1 &= p_L - Lv_L & \epsilon_8 &= \frac{v_{C_1}}{C_1} - \frac{v_{C_3}}{C_3} \\ \epsilon_2 &= p_{C_1} & \epsilon_9 &= q_L - q_{C_2} \\ \epsilon_3 &= p_{C_2} & \epsilon_{10} &= q_{C_1} - q_{C_2} + q_{C_3} \\ \epsilon_4 &= p_{C_3} & \epsilon_{11} &= \nu_L \\ \epsilon_5 &= v_L - v_{C_2} & \epsilon_{12} &= \nu_{C_1} \\ \epsilon_6 &= v_{C_1} - v_{C_2} + v_{C_3} & \epsilon_{13} &= \nu_{C_2} \\ \epsilon_7 &= \frac{q_{C_1}}{C_1} - \frac{q_{C_3}}{C_3} & \epsilon_{14} &= \nu_{C_3}. \end{aligned}$$

We can calculate the matrix  $(\Sigma_{ij}) = (\{\epsilon_i, \epsilon_j\}(q, v, p, \nu))$  as

$$(\Sigma_{ij}) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & -L & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{C_1} & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{C_3} & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 1 \\ 0 & \frac{1}{C_1} & 0 & -\frac{1}{C_3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{C_1} & 0 & -\frac{1}{C_3} & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ L & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & -\frac{1}{C_1} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & \frac{1}{C_3} & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

which is invertible, with inverse  $\Sigma^{ij}$ , then  $\Sigma^{ik}\Sigma_{kj} = \delta_j^i$ . This means that  $x_0 + \widetilde{W}_3$  is a second class submanifold, or equivalently, that all the constraints are second class. In particular,  $x_0 + \widetilde{W}_3$  is a symplectic submanifold, and we will call the symplectic form  $\omega_c$ . Then we have two ways of writing equations of motion on this submanifold.

On one hand, we can write the equations of motion on the symplectic manifold  $x_0 + \widetilde{W}_3$  in the form (2.3),

$$X_{\mathcal{E}|x_0+\widetilde{W}_3} = (\omega_c)^\sharp \left( d(\mathcal{E}|x_0 + \widetilde{W}_3) \right).$$

For this we are going to use the previous parametrization of  $x_0 + \widetilde{W}_3$  with coordinates  $q_L, p_L$ , so  $\omega_c = dq_L \wedge dp_L$ . Since the Lagrangian is given by

$$\mathcal{L}(q, v) = \frac{1}{2}L(v_L)^2 - \frac{1}{2}\frac{(q_{C_1})^2}{C_1} - \frac{1}{2}\frac{(q_{C_2})^2}{C_2} - \frac{1}{2}\frac{(q_{C_3})^2}{C_3},$$

in terms of the parametrization the energy  $\mathcal{E} = pv - \mathcal{L}(q, v)$  restricted to  $x_0 + \widetilde{W}_3$  is

$$\left( \mathcal{E}|x_0 + \widetilde{W}_3 \right) (q_L, p_L) = \frac{1}{2}\frac{p_L^2}{L} + \frac{1}{2}q_L^2 \left( \frac{1}{C_1 + C_3} + \frac{1}{C_2} \right).$$



Then the vector field  $X_{\mathcal{E}|x_0+\widetilde{W}_3}$  is given in coordinates  $q_L, p_L$  by

$$\dot{q}_L = \frac{\partial H}{\partial p_L} = \frac{p_L}{L} \quad (6.43)$$

$$\dot{p}_L = -\frac{\partial H}{\partial q_L} = -q_L \left( \frac{1}{C_1 + C_3} + \frac{1}{C_2} \right), \quad (6.44)$$

with  $H = \mathcal{E}|x_0 + \widetilde{W}_3$ , for short.

We remark that the previous equations can easily be deduced by elementary rules of circuit theory. Namely, one can use the formulas for capacitors in parallel and series and replace the three capacitors by a single one. The resulting circuit is very easy to solve, and the corresponding system is equivalent to equations (6.43) and (6.44). However, we must remark that this simplified system no longer accounts for the currents and voltages on the original capacitors.

On the other hand, we can find the evolution of all the variables  $(q, v, p)$  using the extended energy (3.26) rather than the total energy, since  $x_0 + \widetilde{W}_3$  is symplectic and using the last paragraph of remark 3.30. Namely, we will calculate the vector field  $X = d\mathcal{E}^\sharp + \lambda_{(3)}^j X_{\epsilon_j}$  associated to the extended energy.

We have  $\lambda_{(3)}^i = \Sigma^{ij}\{\mathcal{E}, \epsilon_j\}$ ,  $j = 1, \dots, 14$ , where the column vector  $[\{\mathcal{E}, \epsilon_j\}]$  is

$$[\{\mathcal{E}, \epsilon_j\}] = \left[ 0, -\frac{qC_1}{C_1}, -\frac{qC_2}{C_2}, -\frac{qC_3}{C_3}, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0 \right]^T.$$

For instance, we can calculate  $\dot{p}_L$  and  $\dot{v}_L$  as

$$\begin{aligned} \dot{p}_L &= X(p_L) = (d\mathcal{E})^\sharp(p_L) + \lambda_{(3)}^j X_{\epsilon_j}(p_L) = \{p_L, \mathcal{E}\} + \lambda_{(3)}^j \{p_L, \epsilon_j\} = 0 + \lambda_{(3)}^9 (-1) \\ &= -\frac{qC_1}{C_1} - \frac{qC_2}{C_2} \\ \dot{v}_L &= X(v_L) = (d\mathcal{E})^\sharp(v_L) + \lambda_{(3)}^j X_{\epsilon_j}(v_L) = \{v_L, \mathcal{E}\} + \lambda_{(3)}^j \{v_L, \epsilon_j\} = 0 + \lambda_{(3)}^{11} \\ &= \frac{1}{L} \left( -\frac{qC_1}{C_1} - \frac{qC_2}{C_2} \right) \end{aligned}$$

The complete coordinate expression of  $X$  is

$$\begin{aligned} X &= (\dot{q}_L, \dot{q}_{C_1}, \dot{q}_{C_2}, \dot{q}_{C_3}, \dot{v}_L, \dot{v}_{C_1}, \dot{v}_{C_2}, \dot{v}_{C_3}, \dot{p}_L, \dot{p}_{C_1}, \dot{p}_{C_2}, \dot{p}_{C_3}, \dot{v}_L, \dot{v}_{C_1}, \dot{v}_{C_2}, \dot{v}_{C_3}) \\ &= \left( v_L, v_{C_1}, v_{C_2}, v_{C_3}, \frac{1}{L} \left( -\frac{qC_1}{C_1} - \frac{qC_2}{C_2} \right), \frac{C_1}{C_1 + C_3} \frac{1}{L} \left( -\frac{qC_1}{C_1} - \frac{qC_2}{C_2} \right), \right. \\ &\quad \left. \frac{1}{L} \left( -\frac{qC_1}{C_1} - \frac{qC_2}{C_2} \right), \frac{C_3}{C_1 + C_3} \frac{1}{L} \left( -\frac{qC_1}{C_1} - \frac{qC_2}{C_2} \right), -\frac{qC_1}{C_1} - \frac{qC_2}{C_2}, 0, 0, 0, 0, 0, 0, 0 \right). \end{aligned}$$

Note that this is not a Hamiltonian vector field. However, if we use Theorem 3.28, we can write the equations of motion in terms of the Dirac bracket and the abridged total energy corresponding to the chosen leaf  $x_0 + \widetilde{W}_3$ . In section 7 we will return to LC circuits and show how to write equations of motion on  $M_c$  without specifying

any leaf *a priori*, by using an abridged total energy and the Dirac bracket associated to adapted constraints.

We remark that it is not always the case that there is uniqueness of solution. In fact, several branches of the circuit may be present with no capacitors or inductors, as a limit case. In the next section we show how to write a Poisson bracket description of the solution.

## 7 An extension of the Dirac theory of constraints

In sections 2 and 3, under precise regularity conditions, we studied several geometric and algebraic objects related to the Dirac theory of constraints like primary and final constraints, first class and second class constraints, Dirac brackets, first class and second class constraint submanifolds.

The notions of first class and second class depend only on the final constraint submanifold  $M_c$  and they do not depend on the primary constraint submanifold or the Hamiltonian, which are needed only to write equations of motion. Then, as we have already said at the beginning of section 3.3, in order to study those notions one can start with an abstract situation, given by a symplectic manifold  $(P, \Omega)$  and an arbitrary submanifold  $S \subseteq P$ , to be thought of as the final constraint submanifold.

In this section, motivated by integrable nonholonomic mechanics and LC circuits theory, we extend those studies by replacing  $S$  by a submanifold  $\mathbf{S} \subseteq P$  regularly foliated by submanifolds  $S$ . This means that  $\mathbf{S} = \cup_{S \in \mathfrak{S}} S$  and the map  $\mathbf{S} \rightarrow \mathfrak{S}$  given by  $x \mapsto S$  iff  $x \in S$  is a submersion. To study the dynamics, one should consider a primary *regularly foliated* constraint submanifold  $\mathbf{S}'$  whose leaves  $S'$  satisfy the condition  $S = S' \cap \mathbf{S}$ , and an energy  $\mathcal{E}: P \rightarrow \mathbb{R}$ .

The meaning of  $\mathbf{S}$  can be interpreted in connection with the constraint algorithm CAD for a Dirac dynamical system (1.1) on the manifold  $M$  in the case in which the Dirac structure  $D$  is integrable, which gives a foliation of  $M$ . We should compare  $\mathbf{S}$  with the foliated submanifold  $M_c$  appearing in remark 5.3, while  $\mathbf{S}'$  should be compared with  $M$  as a submanifold of  $P$ .

The results in section 3.4 can be extended for a certain kind of Dirac dynamical systems. Namely, let  $\mathbf{S} \subseteq \mathbf{S}' \subseteq P$  as before. Consider each leaf  $S'_C$  endowed with the presymplectic form  $\omega_C$  obtained as the pull-back of  $\Omega$ . Assume that the distribution  $\ker \omega_C$  is regular and its dimension does not depend on  $C$ . Define the Dirac structure  $D$  on  $\mathbf{S}'$  by declaring that its presymplectic leaves are  $(S'_C, \omega_C)$ . We will consider the Dirac dynamical system (1.1) for the case  $M = \mathbf{S}'$ , that is,

$$(x, \dot{x}) \oplus d\mathcal{E}(x) \in D_x. \quad (7.1)$$

The results in section 3.4 would correspond to the case in which the foliation of  $\mathbf{S}$  consists of a single leaf.

## 7.1 Dirac brackets adapted to foliated constraint submanifolds

**Description of primary and final foliated constraint submanifolds.** Let  $(P, \Omega)$  be a symplectic manifold. We are going to describe the *primary foliated constraint submanifold*  $\mathbf{S}'$  and the *final foliated constraint submanifold*  $\mathbf{S}$  of  $P$ . We will assume that there is a one to one correspondence between the foliations such that for each leaf  $S' \subseteq \mathbf{S}'$  the corresponding leaf of  $\mathbf{S}$  is  $S = S' \cap \mathbf{S}$ . A motivation for this assumption comes from mechanics, where an energy on  $P$  is given. In that context, each leaf  $S \subseteq \mathbf{S}$  is the final constraint submanifold corresponding to the primary constraint submanifold  $S' \subseteq \mathbf{S}'$  after applying a constraint algorithm. These leaves will be parametrized by a certain vector  $C$ , and we will denote them by  $S'_C$  and  $S_C$  respectively. We are going to work on a suitable neighborhood  $\mathcal{U}$  of  $\mathbf{S}'$  in  $P$ , where all these submanifolds can be defined regularly by equations, as we will describe next.

For convenience, we will work with a more general situation where  $\mathbf{S}'$  is a member of a *family* of primary foliated constraint submanifolds parameterized by a certain vector  $C_{(1,a')}$  and  $\mathbf{S}$  is a member of a corresponding family of final foliated constraint submanifolds parameterized by a certain vector  $C_{(1,a)}$ . For  $C_{(1,a')} = 0$  and  $C_{(1,a)} = 0$  we obtain  $\mathbf{S}'$  and  $\mathbf{S}$ , respectively.

Let

$$C = (C_1, \dots, C_{a'}, C_{a'+1}, \dots, C_a, C_{a+1}, \dots, C_b) \equiv (C_{(1,a')}, C_{(a'+1,a)}, C_{(a+1,b)}) \quad (7.2)$$

be a generic point of  $\mathbb{R}^b \equiv \mathbb{R}^{a'} \times \mathbb{R}^{a-a'} \times \mathbb{R}^{b-a}$ , where  $a' \leq a \leq b$ . Denote  $C_{(1,a)} = (C_1, \dots, C_a)$ . Then we assume the following definitions by equations (regularly on  $\mathcal{U}$ ) of the submanifolds  $\mathbf{S}^{C_{(1,a'')}}$  and  $\mathbf{S}^{C_{(1,a)}}$  and their foliations by leaves  $S'_C$  and  $S_C$  respectively, which satisfy  $S_C = S'_C \cap \mathbf{S}^{C_{(1,a)}}$ , namely,

$$\phi_i = C_i, i = 1, \dots, a', \text{ defines } \mathbf{S}^{C_{(1,a')}} \quad (7.3)$$

$$\phi_i = C_i, i = 1, \dots, a', a' + 1, \dots, a, \text{ defines } \mathbf{S}^{C_{(1,a)}} \quad (7.4)$$

$$\phi_i = C_i, i = 1, \dots, a', a + 1, \dots, b, \text{ defines } S'_C \quad (7.5)$$

$$\phi_i = C_i, i = 1, \dots, a, a + 1, \dots, b, \text{ defines } S_C \quad (7.6)$$

Since for each primary constraint submanifold  $S'_C$  there is only one final constraint submanifold  $S_C$ , then  $C_{(a'+1,a)}$  must be a function of  $C_{(1,a')}$ , and without loss of generality we will assume that this function maps 0 to 0. Then for fixed  $C_{(1,a')}$ ,  $C_{(a'+1,a)}$  is also fixed and one has, for each choice of  $C_{(a+1,b)} = (C_{a+1}, \dots, C_b)$ , one leaf  $S'_C$  of  $\mathbf{S}^{C_{(1,a'')}}$  and the corresponding leaf  $S_C = S'_C \cap \mathbf{S}^{C_{(1,a)}}$  of  $\mathbf{S}^{C_{(1,a)}}$ .

**Assumption 7.1.** *From now we assume that our foliated submanifolds  $\mathbf{S}'$  and  $\mathbf{S}$  are  $\mathbf{S}^{C_{(1,a'')}}$  and  $\mathbf{S}^{C_{(1,a)}}$  for the choice  $C_{(1,a')} = 0$  and  $C_{(1,a)} = 0$ , respectively, while the leaves  $S'$  and  $S$  are the leaves  $S'_C$  and  $S_C$  respectively, for the choice  $C = (0, 0, C_{(a+1,b)})$ , and  $C_{(a+1,b)}$  varying arbitrarily, that is,  $C \in \{0\} \times \{0\} \times \mathbb{R}^{b-a}$  (see Figure 7.1).*

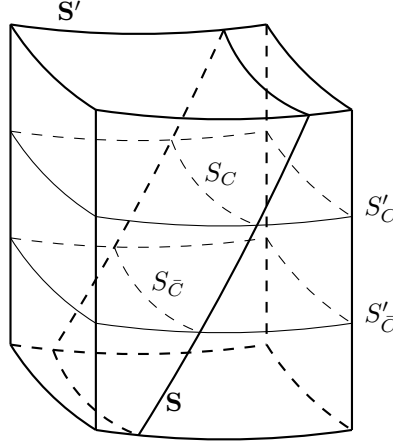


Figure 7.1:  $\mathbf{S}'$  is foliated by the submanifolds  $\{S'_C\}$ , while  $\mathbf{S}$  is foliated by  $\{S_C\}$ . Here we depict the leaves for two values  $C$  and  $\bar{C}$ .

*Remark 7.2.* The case considered by Dirac, studied in section 3, corresponds to  $a = b$ . It is in this sense that our theory extends Dirac's. In this case, the primary constraint  $\mathbf{S}'$  and the final constraint  $\mathbf{S}$  have only one leaf. More precisely, this particular case has been studied in Lemma 3.23, with a different notation. Our extension takes care of those Dirac dynamical systems coming for instance from integrable nonholonomic systems or LC circuit theory.

**Equations of motion as a collection of Hamilton's equations.** Let us adopt the point of view that the Dirac dynamical system (7.1) is a collection of Gotay-Nester systems, one for each leaf, namely

$$\Omega(x)(\dot{x}, \delta x) = d\mathcal{E}(x)(\delta x), \quad (7.7)$$

where  $(x, \dot{x}) \in T_x S_C$ , for all  $\delta x \in T_x S'_C$ .

Then one can apply the theory developed in section 3.4 leaf by leaf. For each  $C = (0, 0, C_{(a+1, b)})$ , the total Hamiltonian defined in (3.19) should be replaced by the total energy depending on  $C$ , by definition,

$$\mathcal{E}_{C,T} = \mathcal{E} + \lambda^i(\phi_i - C_i), \quad \text{sum over } i = 1, \dots, a', a+1, \dots, b$$

and, for each such  $C$ , the  $\lambda^i \in C^\infty(\mathcal{U})$ ,  $i = 1, \dots, a', a+1, \dots, b$  must satisfy the conditions  $\{\mathcal{E}_{C,T}, \phi_j - C_j\}(x) = 0$ ,  $j = 1, \dots, b$ , for each  $x \in S_C$  or, equivalently,

$$\{\mathcal{E}, \phi_j\}(x) + \lambda^i \{\phi_i, \phi_j\}(x) = 0, \quad j = 1, \dots, b,$$

for each  $x \in S_C$ . This shows in particular the interesting fact that the set of solutions, which is an affine space, depends on  $x$  but not on  $C$ . We assume, as part of our regularity conditions, that the solutions  $(\lambda^1, \dots, \lambda^{a'}, \lambda^{a+1}, \dots, \lambda^b)$  form a nonempty affine bundle  $\Lambda \rightarrow \mathbf{S}$ . Since we are working locally, by shrinking  $\mathcal{U}$  conveniently we can assume that this bundle is trivial. Furthermore, we extend this

bundle arbitrarily to a trivial affine bundle over  $\mathcal{U}$ , which we continue to denote  $\Lambda \rightarrow \mathcal{U}$ . Note that for each  $C$ , the total energy can be seen as a function  $\mathcal{E}_{C,T}: \Lambda \rightarrow \mathbb{R}$ .

For each section  $\tilde{\lambda}$  of  $\Lambda$  one can define  $\mathcal{E}_{C,\tilde{\lambda},T} = \tilde{\lambda}^* \mathcal{E}_{C,T}: \mathcal{U} \rightarrow \mathbb{R}$ , that is,  $\mathcal{E}_{C,\tilde{\lambda},T}(x) = \mathcal{E}(x) + \tilde{\lambda}^i(x)(\phi_i(x) - C_i)$ . One has an equation of motion on  $\mathcal{U}$ ,

$$X_{\mathcal{E}_{C,\tilde{\lambda},T}}(x) = (d\mathcal{E}_{C,\tilde{\lambda},T})^\#(x), \quad (7.8)$$

which should be interpreted properly, as follows. In the definition of  $\mathcal{E}_{C,T}$ ,  $x$  and  $C$  are independent variables. For a given  $x_0 \in \mathcal{U}$ , let  $C_0 = (\phi_1(x_0), \dots, \phi_b(x_0))$ ; then  $\mathcal{E}_{C_0,\tilde{\lambda},T}(x)$  is a function of  $x$  and

$$X_{\mathcal{E}_{C_0,\tilde{\lambda},T}}(x_0) = (d\mathcal{E}_{C_0,\tilde{\lambda},T})^\#(x_0) = d\mathcal{E}^\#(x_0) + \tilde{\lambda}^i(x_0)d\phi_i^\#(x_0).$$

This means that equation (7.8) should not be naively interpreted as Hamilton's equation, and therefore the algorithms devised by Dirac and Gotay-Nester cannot be generalized in a direct way for the foliated case. In fact, even though the variable  $C$  is a function of  $x$ , it should be considered a constant in order to calculate the vector field at the point  $x$ . Note that, by construction, the condition  $X_{\mathcal{E}_{C_0,\tilde{\lambda},T}}(x_0) \in T_{x_0}S_{C_0}$  is satisfied, which means that the leaves  $S_{C_0}$  of  $\mathbf{S}$  are preserved by the motion, as expected.

*Remark 7.3.* For any given solution  $x(t)$  of equations of motion (7.7) there exist uniquely determined  $\lambda_t^i$ ,  $i = 1, \dots, a', a + 1, \dots, b$  such that

$$\dot{x}(t) = d\mathcal{E}^\#(x(t)) + \lambda_t^i d\phi_i^\#(x(t)).$$

On the other hand, for any given time-dependent section  $\tilde{\lambda}_t^i(x)$ ,  $i = 1, \dots, a', a + 1, \dots, b$  of  $\Lambda$  one has a time-dependent vector field

$$d\mathcal{E}^\#(x) + \tilde{\lambda}_t^i(x)d\phi_i^\#(x) \quad (7.9)$$

whose integral curves are solutions to the equations of motion. One can show that under the strong regularity conditions that we assume in this paper, all the solutions of the equations of motion can be represented in this way, at least locally.

Summarizing, the vector fields (7.9) obtained from (7.8) are not in general Hamiltonian vector fields with respect to the canonical bracket, as we have indicated before. Now we will see how to recover the Hamiltonian character of the equation of motion. We are going to write it as Hamilton's equation in terms of a certain Poisson bracket whose symplectic leaves are isomorphic to  $\mathcal{U}$ .

**Equations of motion in terms of a Poisson bracket.** Let us consider an extended Poisson manifold, namely the product Poisson manifold  $(\mathcal{U} \times \{0\} \times \{0\} \times \mathbb{R}^{b-a}, \{, \}^{(b)})$  with the canonical Poisson bracket  $(\Omega)^{-1}$  on  $\mathcal{U}$  and the 0 Poisson bracket on  $\{0\} \times \{0\} \times \mathbb{R}^{b-a}$ . Let  $\mathcal{E}_T: \Lambda \times \{0\} \times \{0\} \times \mathbb{R}^{b-a} \rightarrow \mathbb{R}$  be defined by  $\mathcal{E}_T(\lambda, C) := \mathcal{E}_{C,T}(\lambda)$ . Given a section  $\tilde{\lambda}$  of  $\Lambda$  one can define an energy function

$\mathcal{E}_{\tilde{\lambda},T}: \mathcal{U} \times \{0\} \times \{0\} \times \mathbb{R}^{b-a} \rightarrow \mathbb{R}$  as  $\mathcal{E}_{\tilde{\lambda},T}(x, C) := \mathcal{E}_{C,T}(\tilde{\lambda}(x)) = \mathcal{E}_{C,\tilde{\lambda},T}(x)$ . Consider its corresponding Hamiltonian vector field

$$X_{\mathcal{E}_{\tilde{\lambda},T}}(x, C) = (X_{\mathcal{E}_{C,\tilde{\lambda},T}}(x), 0),$$

on  $\mathcal{U} \times \{0\} \times \{0\} \times \mathbb{R}^{b-a}$  with respect to the Poisson bracket  $\{, \}^{(b)}$ , where  $X_{\mathcal{E}_{C,\tilde{\lambda},T}}(x)$  is defined in (7.8). We remark that even though the first component cannot be interpreted in general as a Hamiltonian vector field on  $\mathcal{U}$ , as we have observed before,  $X_{\mathcal{E}_{\tilde{\lambda},T}}$  is Hamiltonian on the extended Poisson manifold.

Consider the function

$$\mathcal{C}_{(0,0)}(x) = (0, 0, \phi_{a+1}(x), \dots, \phi_b(x)) \in \mathbb{R}^{a'} \times \mathbb{R}^{a-a'} \times \mathbb{R}^{b-a}$$

defined on  $\mathcal{U}$ . Clearly,  $\text{graph}(\mathcal{C}_{(0,0)})$  is a submanifold of  $\mathcal{U} \times \{0\} \times \{0\} \times \mathbb{R}^{b-a}$  diffeomorphic to  $\mathcal{U}$ .

Since  $X_{\mathcal{E}_{C,\tilde{\lambda},T}}$  preserves  $S_C$  for each  $C$ , it is straightforward to prove that  $X_{\mathcal{E}_{\tilde{\lambda},T}}$  preserves  $\text{graph}(\mathcal{C}_{(0,0)}|_{S_C})$ , which for each  $C$  is a copy of the level set  $\mathcal{C}_{(0,0)}^{-1}(C)$  embedded in  $\mathcal{U} \times \{0\} \times \{0\} \times \mathbb{R}^{b-a}$ , and therefore  $X_{\mathcal{E}_{\tilde{\lambda},T}}$  preserves  $\text{graph}(\mathcal{C}_{(0,0)}|\mathbf{S})$ , which is a copy of  $\mathbf{S}$ .

Consider the evolution equation for a quantity  $F$  on  $\mathcal{U} \times \{0\} \times \{0\} \times \mathbb{R}^{b-a}$  given by

$$\dot{F}(x, C) = \{F, \mathcal{E}_{\tilde{\lambda},T}\}^{(b)}(x, C).$$

Note that this is an equation in Poisson form on  $\mathcal{U} \times \{0\} \times \{0\} \times \mathbb{R}^{b-a}$ , not on  $\mathbf{S} \times \{0\} \times \{0\} \times \mathbb{R}^{b-a}$ . Note that for  $a = b$  and  $P = T^*Q$ , then  $\mathbf{S} = S$  and we recover the situation in Dirac's theory, where Hamilton's equations of motion are written in a neighborhood of the final constraint submanifold in  $T^*Q$  rather than on the constraint submanifold itself.

**Local equations of motion in Poisson form with respect to the Dirac bracket.** As we have remarked above, the equations of motion in terms of the canonical bracket for the foliated case are not naive extensions of the Dirac or Gotay-Nester procedures. A similar situation occurs with the equation  $\dot{g} \approx [g, H_T]^*$  in Dirac [1964], page 42. As we will show next, we can extend this equation to the foliated case by using adapted constraints and the abridged total energy.

In the definition of  $S'_C$  and  $S_C$  take  $C = 0$ . Then we can apply the procedure of section 3.4 with  $S = S_0$  and  $S' = S'_0$ , and obtain second class constraints  $\chi_1, \dots, \chi_{2s}$  among  $\phi_i$ ,  $i = 1, \dots, b$  adapted to  $S'_0$  and also we can choose some primary constraints among  $\phi_i$ ,  $i = 1, \dots, a'$ ,  $a + 1, \dots, b$ , say w.l.o.g.  $\phi_k$ ,  $k = 1, \dots, a' - s'_A$ ,  $a + 1, \dots, b - s'_B$ . Observe that here  $s' = s'_A + s'_B$  has the same meaning as in section 3.4, this time considering the constraints  $\phi_1 = 0, \dots, \phi_b = 0$ . Then the equations of motion on  $S_0$  can be written in the form (3.25).

Now, for any small enough  $C$ , which defines  $S_C$ , one can readily see that  $\chi_1 - B_1, \dots, \chi_{2s} - B_{2s}$  are second class constraints adapted to  $S'_C$  where  $B_1, \dots, B_{2s}$  are the components of  $C$  corresponding to the choice  $\chi_1, \dots, \chi_{2s}$ . This implies

immediately that the Dirac bracket for  $S_C$  does not depend on  $C$ . On the other hand, each  $\phi_k - C_k$ ,  $k = 1, \dots, a' - s'_A, a + 1, \dots, b - s'_B$ , is a primary constraint for  $S'_C$ .

Using the previous facts and the fact that the  $\lambda^k$  are arbitrary (even time-dependent) parameters, as it happens with the  $\lambda^i$  in equation (3.25), one can conclude that the equation of motion on each  $S_C$  is given by

$$\begin{aligned}\dot{F} &= \{F, \mathcal{E}\}_{(\chi)} + \lambda^k \{F, \phi_k\}_{(\chi)} \\ &= \{F, \mathcal{E} + \lambda^k \phi_k\}_{(\chi)},\end{aligned}\tag{7.10}$$

sum over  $k = 1, \dots, a' - s'_A, a + 1, \dots, b - s'_B$ .

This means in particular that each  $S_C$  is preserved by the motion, or that the Hamiltonian vector fields

$$X_{(\chi), \mathcal{E} + \lambda^k \phi_k},\tag{7.11}$$

defined on  $U$ , are tangent to the  $S_C$ .

Note that  $\mathcal{E} + \lambda^k \phi_k$  is the abridged total energy  $\mathcal{E}_{AT}$  for  $S_0$  and  $S'_0$  as defined in (3.23).

In conclusion, using adapted constraints, the abridged total energy and the Dirac bracket yields a very simple procedure for writing the equations of motion, because one only needs to write the abridged total energy for one leaf, say  $C = 0$ , and (7.10) gives the equation of motion for all nearby leaves.

We have proven the following theorem, which extends Theorem 3.28.

**Theorem 7.4.** *Let  $(P, \Omega)$  be a symplectic manifold and let  $\mathbf{S}' \supseteq \mathbf{S}$  be given primary and final foliated constraint submanifolds, with leaves  $S'_C$  and  $S_C = S'_C \cap \mathbf{S}$ , respectively, as described by (7.2)–(7.6). Assume that all the hypotheses of Theorem 3.28 are satisfied for the submanifolds  $S'_0 \supseteq S_0$ . In addition, suppose that the number  $s'$  appearing in Assumption 3.26 is the same for all  $S_C$  for  $C$  close enough to 0. From the validity of Assumption 3.27 for  $S'_0$ , it is immediate to see that the number  $2s$  is also the same for all  $S'_C$  for  $C$  close enough to 0. Choose second class constraints  $(\chi'_1, \dots, \chi'_{s'}, \chi''_{s'+1}, \dots, \chi''_{2s})$  adapted to  $S'_0$  and also  $\phi_k$ ,  $k = 1, \dots, a' - s'_A, a + 1, \dots, b - s'_B$ , as in Theorem 3.28.*

*Let  $\mathcal{E}: P \rightarrow \mathbb{R}$  be an energy function, and consider the Dirac dynamical system (7.1). Define the abridged total energy  $\mathcal{E}_{AT} = \mathcal{E} + \lambda^k \phi_k$  for  $S_0$  and  $S'_0$  as in Theorem 3.28. Then, each  $x_0 \in \mathbf{S}$  has an open neighborhood  $U$  such that the vector field  $X_{(\chi), \mathcal{E}_{AT}}$  (equation (7.11)) defined on  $U$ , when restricted to the final foliated constraint submanifold  $\mathbf{S} \cap U$ , represents equations of motion of the Dirac dynamical system on  $\mathbf{S} \cap U$ . Moreover, the evolution of the system preserves the leaves  $S_C \cap U$ . In addition, equation (7.10) gives equations of motion on each  $S_C \cap U$ . Since  $X_{(\chi), \mathcal{E}_{AT}}$  is tangent to  $S_C \cap U$ , then the evolution of a function  $f$  on  $S_C \cap U$  is given by (7.10) for any  $F$  such that  $f = F|_{S_C \cap U}$ . Also, each  $S_C \cap U$  is contained in a unique symplectic leaf of the Dirac bracket, defined by  $\chi'_1 = B_1, \dots, \chi'_{s'} = B_{s'}, \chi''_{s'+1} = B_{s'+1}, \dots, \chi''_{2s} = B_{2s}$ .*

**Remark 7.5.** Dirac [1964, page 42], wrote the equation

$$\dot{g} \approx \{g, H_T\}^*$$

which represents the dynamics on the final constraint submanifold, in terms of the total Hamiltonian and the Dirac bracket. This equation does not gives the correct dynamics for the case of foliated constraint submanifolds. As we have seen, for the foliated case, one has, instead a similar equations of motion (7.10) in terms of the Dirac bracket and the abridged total energy  $\mathcal{E}_{AT}$ .

**The concrete example of an LC circuit of section 6 revisited.** Consider the symplectic leaf  $S$  of  $M_3$  defined parametrically by equations (6.38)–(6.42). The parameter space carries the canonical symplectic form  $dq_L \wedge dp_L$ , which coincides with the pullback of the canonical symplectic form on  $T^*TQ$  to the parameter space. Since the matrix  $\Sigma$  is invertible, all constraints are second class constraints. The number  $b - s'_B$  in equation (7.10) is 0, then the abridged total energy is simply the energy  $\mathcal{E}$ .

Then it is easy to calculate the Dirac bracket as the canonical bracket on the parameter space of the functions (6.38)–(6.42). Let  $y = (q, v, p, \nu)$ , then the matrix of Dirac brackets of these variables can be written in block form as

$$(\{y^i, y^j\}_\chi) = \begin{bmatrix} 0_{4 \times 4} & A & 0_{4 \times 7} \\ -A^T & 0_{5 \times 5} & 0_{5 \times 7} \\ 0_{7 \times 4} & 0_{7 \times 5} & 0_{7 \times 7} \end{bmatrix}$$

where

$$A = \begin{bmatrix} \frac{1}{L} & \frac{C_1}{(C_1+C_3)L} & \frac{1}{L} & \frac{C_3}{(C_1+C_3)L} & 1 \\ \frac{C_1}{(C_1+C_3)L} & \frac{C_1^2}{(C_1+C_3)^2L} & \frac{C_1}{(C_1+C_3)L} & \frac{C_1C_3}{(C_1+C_3)^2L} & \frac{C_1}{(C_1+C_3)} \\ \frac{1}{L} & \frac{C_1}{(C_1+C_3)L} & \frac{1}{L} & \frac{C_3}{(C_1+C_3)L} & 1 \\ \frac{C_3}{(C_1+C_3)L} & \frac{C_1C_3}{(C_1+C_3)^2L} & \frac{C_3}{(C_1+C_3)L} & \frac{C_3^2}{(C_1+C_3)^2L} & \frac{C_3}{(C_1+C_3)} \end{bmatrix}.$$

The evolution equations for the vector

$$(y^1, \dots, y^{16}) = (q_L, q_{C_1}, q_{C_2}, q_{C_3}, v_L, v_{C_1}, v_{C_2}, v_{C_3}, p_L, p_{C_1}, p_{C_2}, p_{C_3}, \nu_L, \nu_{C_1}, \nu_{C_2}, \nu_{C_3})$$

with initial condition belonging to  $M_3$ , that is, satisfying (6.36):  $p_L = Lv_L$ ,  $p_{C_1} = p_{C_2} = p_{C_3} = 0$ ,  $v_L = v_{C_2}$ ,  $v_{C_1} = v_{C_2} - v_{C_3}$ ,  $\frac{q_{C_1}}{C_1} = \frac{q_{C_3}}{C_3}$ ,  $\frac{v_{C_1}}{C_1} = \frac{v_{C_3}}{C_3}$ , is given by the product of the matrix  $(\{y^i, y^j\}_\chi)$  by the vector

$$\left( \frac{\partial \mathcal{E}}{\partial y^1}, \dots, \frac{\partial \mathcal{E}}{\partial y^{16}} \right) = \left( 0, -\frac{q_{C_1}}{C_1}, -\frac{q_{C_2}}{C_2}, -\frac{q_{C_3}}{C_3}, p_L - v_L, p_{C_1}, p_{C_2}, p_{C_3}, v_L, v_{C_1}, v_{C_2}, v_{C_3}, 0, 0, 0, 0 \right).$$

## 8 Conclusions and future work

We have shown that Dirac's work on constrained systems can be extended for cases where the primary constraint has a given foliation. This extends the applicability of



the theory for Dirac dynamical systems, like LC circuits where the constraints may come from singularities of the Lagrangian and, besides, from Kirchhoff's Current Law. Throughout the paper we combine ideas of an algebraic character from Dirac's theory with some more geometrically inspired ideas from Gotay-Nester work. In particular, we use Dirac brackets adapted to the foliation as well as a Constraint Algorithm for Dirac dynamical systems (CAD), which is an extension of the Gotay-Nester algorithm. The results were proven locally and under regularity conditions. It is our purpose to study in the future the globalization of the results of the paper as well as the singular cases, using IDE techniques, with applications.

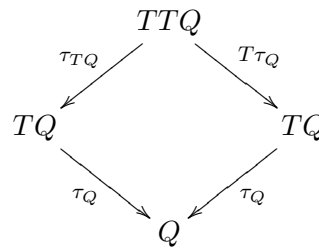
## 9 Acknowledgments

This paper has been inspired by Jerry Marsden's invaluable participation in part of it. We thank the following institutions for providing us with means to work on this paper: Universidad Nacional del Sur (projects PGI 24/L075 and PGI 24/ZL06); Universidad Nacional de la Plata; Agencia Nacional de Promoción Científica y Tecnológica, Argentina (projects PICT 2006-2219 and PICT 2010-2746); CONICET, Argentina (projects PIP 2010-2012 11220090101018 and X553 Ecuaciones Diferenciales Implícitas); European Community, FP7 (project IRSES "GEOMECH" 246981).

## A Appendix

**Lemma A.1.** *There is a canonical inclusion  $\varphi: TQ \oplus T^*Q \rightarrow T^*TQ$ . In addition, consider the canonical two-forms  $\omega_{T^*Q}$  and  $\omega_{T^*TQ}$  on  $T^*Q$  and  $T^*TQ$  respectively, the canonical projection  $\text{pr}_{T^*Q}: TQ \oplus T^*Q \rightarrow T^*Q$ , and define the presymplectic two-form  $\omega = \text{pr}_{T^*Q}^* \omega_{T^*Q}$  on  $TQ \oplus T^*Q$ . Then the inclusion preserves the corresponding two-forms, that is,  $\omega = \varphi^* \omega_{T^*TQ}$ .*

**Proof.** If  $\tau_Q: TQ \rightarrow Q$  and  $\tau_{TQ}: TTQ \rightarrow TQ$  are the tangent projections, we can consider the dual tangent rhombic



Define  $\varphi: TQ \oplus T^*Q \rightarrow T^*TQ$  by  $\varphi(v_q \oplus \alpha_q) \in T_{v_q}^*TQ$ ,

$$\varphi(v_q \oplus \alpha_q) \cdot w_{v_q} = \alpha_q \cdot T\tau_Q(w_{v_q}),$$

for  $w_{v_q} \in T_{v_q}TQ$ . Here  $v_q \oplus \alpha_q$  denotes an element in the Pontryagin bundle over

the point  $q \in Q$ . Note that the following diagram commutes.

$$\begin{array}{ccc} TQ \oplus T^*Q & \xrightarrow{\varphi} & T^*TQ \\ \text{pr}_{TQ} \downarrow & & \downarrow \pi_{TQ} \\ TQ & \xrightarrow{\text{Id}_{TQ}} & TQ \end{array}$$

Let us see that  $\varphi$  is an injective vector bundle map from the bundle  $\text{pr}_{TQ}: TQ \oplus T^*Q \rightarrow TQ$  to the cotangent bundle  $\pi_{TQ}: T^*TQ \rightarrow TQ$ , over the identity of  $TQ$ . The last part of this assertion follows from the commutative diagram above.

First, if  $\varphi(v_q \oplus \alpha_q) = \varphi(v'_{q'} \oplus \alpha'_{q'})$  then both sides are in the same fiber  $T^*_{v_q} TQ = T^*_{v'_{q'}} TQ$ , so  $v_q = v'_{q'}$ . Also, for all  $w_{v_q} \in T^*_{v_q} TQ$  we have

$$\varphi(v_q \oplus \alpha_q) \cdot w_{v_q} = \varphi(v_q \oplus \alpha'_q) \cdot w_{v_q}$$

or

$$\alpha_q \cdot T\tau_Q(w_{v_q}) = \alpha'_q \cdot T\tau_Q(w_{v_q}).$$

Since  $\tau_Q: TQ \rightarrow Q$  is a submersion, we have  $\alpha_q = \alpha'_q$  and  $\varphi$  is injective.

Second,  $\varphi$  is linear on each fiber, since

$$\varphi(v_q \oplus (\alpha_q + \lambda \alpha'_q)) \cdot w_{v_q} = (\alpha_q + \lambda \alpha'_q) \cdot T\tau_Q(w_{v_q}) = \varphi(v_q \oplus \alpha_q) \cdot w_{v_q} + \lambda \varphi(v_q \oplus \alpha'_q) \cdot w_{v_q}.$$

For the second part of the lemma, let us recall the definition of the canonical one-form on  $\theta_{T^*Q} \in \Omega^1(T^*Q)$ . For  $\alpha_q \in T^*Q$ ,  $\theta_{T^*Q}(\alpha_q)$  is an element of  $T^*_{\alpha_q} T^*Q$  such that for  $w_{\alpha_q} \in T_{\alpha_q} T^*Q$ ,

$$\theta_{T^*Q}(\alpha_q) \cdot w_{\alpha_q} = \alpha_q(T\pi_Q(w_{\alpha_q})),$$

where  $\pi_Q: T^*Q \rightarrow Q$  is the cotangent bundle projection.

With a similar notation, the canonical one-form  $\theta_{T^*TQ} \in \Omega^1(T^*TQ)$  is given by

$$\theta_{T^*TQ}(\alpha_{v_q}) \cdot w_{\alpha_{v_q}} = \alpha_{v_q}(T\pi_{TQ}(w_{\alpha_{v_q}})).$$

Pulling back these forms to the Pontryagin bundle by  $\varphi$  and the projection  $\text{pr}_{T^*Q}: TQ \oplus T^*Q \rightarrow T^*Q$ , we obtain the same one-form. Indeed, for  $w \in T_{v_q \oplus \alpha_q}(TQ \oplus T^*Q)$ , we get on one hand

$$(\text{pr}_{T^*Q}^* \theta_{T^*Q})(v_q \oplus \alpha_q) \cdot w = \theta_{T^*Q}(\alpha_q) \cdot T\text{pr}_{T^*Q}(w) = \alpha_q \cdot T(\pi_Q \circ \text{pr}_{T^*Q})(w),$$

and on the other hand

$$\begin{aligned} (\varphi^* \theta_{T^*TQ})(v_q \oplus \alpha_q) \cdot w &= \theta_{T^*TQ}(\varphi(v_q \oplus \alpha_q)) \cdot T\varphi(w) = \\ &= \varphi(v_q \oplus \alpha_q) \cdot T\pi_{TQ}(T\varphi(w)) = \alpha_q \cdot T(\tau_Q \circ \pi_{TQ} \circ \varphi)(w). \end{aligned}$$

However, the following diagram commutes

$$\begin{array}{ccc} TQ \oplus T^*Q & \xrightarrow{\varphi} & T^*TQ \\ \text{pr}_{T^*Q} \downarrow & \searrow \text{pr}_{TQ} & \downarrow \pi_{TQ} \\ T^*Q & & TQ \\ & \searrow \pi_Q & \swarrow \tau_Q \\ & Q & \end{array}$$

so  $\pi_Q \circ \text{pr}_{T^*Q} = \tau_Q \circ \pi_{TQ} \circ \varphi$  and therefore  $\text{pr}_{T^*Q}^* \theta_{T^*Q} = \varphi^* \theta_{T^*TQ}$ . Taking (minus) the differential of this identity, we obtain  $\omega = \varphi^* \omega_{T^*TQ}$ . ■

In local coordinates, if we denote the elements of  $TQ \oplus T^*Q$  and  $T^*TQ$  by  $(q, v, p)$  and  $(q, v, p, \nu)$  respectively, then  $\varphi(q, v, p) = (q, v, p, 0)$ .

## References

- Ralph Abraham and Jerrold E. Marsden. *Foundations of mechanics*. Benjamin/Cummings Publishing Co. Inc. Advanced Book Program, Reading, Mass., 1978. ISBN 0-8053-0102-X. Second edition, revised and enlarged, with the assistance of Tudor Ratiu and Richard Cushman.
- Vladimir I. Arnol'd. *Mathematical methods of classical mechanics*, volume 60 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1989. ISBN 0-387-96890-3. Translated from the Russian by K. Vogtmann and A. Weinstein.
- P. Balseiro, M. de León, J. C. Marrero, and D. Martín de Diego. The ubiquity of the symplectic Hamiltonian equations in mechanics. *J. Geom. Mech.*, 1(1):1–34, 2009. ISSN 1941-4889. doi: 10.3934/jgm.2009.1.1. URL <http://dx.doi.org/10.3934/jgm.2009.1.1>.
- María Barbero-Liñán and Miguel C. Muñoz-Lecanda. Constraint algorithm for extremals in optimal control problems. *Int. J. Geom. Methods Mod. Phys.*, 6(7):1221–1233, 2009. ISSN 0219-8878. doi: 10.1142/S0219887809004193. URL <http://dx.doi.org/10.1142/S0219887809004193>.
- Guido Blankenstein and Tudor S. Ratiu. Singular reduction of implicit Hamiltonian systems. *Rep. Math. Phys.*, 53(2):211–260, 2004. ISSN 0034-4877. doi: 10.1016/S0034-4877(04)90013-4. URL [http://dx.doi.org/10.1016/S0034-4877\(04\)90013-4](http://dx.doi.org/10.1016/S0034-4877(04)90013-4).
- Anthony M. Bloch. *Nonholonomic mechanics and control*, volume 24 of *Interdisciplinary Applied Mathematics*. Springer-Verlag, New York, 2003. ISBN 0-387-95535-6. With the collaboration of J. Baillieul, P. Crouch and J. Marsden, With scientific input from P. S. Krishnaprasad, R. M. Murray and D. Zenkov, Systems and Control.
- Anthony M. Bloch and Peter E. Crouch. Representations of Dirac structures on vector spaces and nonlinear L-C circuits. In *Differential geometry and control (Boulder, CO, 1997)*, volume 64 of *Proc. Sympos. Pure Math.*, pages 103–117. Amer. Math. Soc., Providence, RI, 1999.
- A. V. Borisov and I. S. Mamaev. On the history of the development of the non-holonomic dynamics. *Regul. Chaotic Dyn.*, 7(1):43–47, 2002. ISSN 1560-3547. doi: 10.1070/RD2002v007n01ABEH000194. URL <http://dx.doi.org/10.1070/RD2002v007n01ABEH000194>.

- Henrique Bursztyn. Dirac structures and applications. La reconquête de la dynamique par la géométrie après Lagrange. Conference at IHES, March 2010. URL [http://www.ihes.fr/document?id=2396&id\\_attribute=48](http://www.ihes.fr/document?id=2396&id_attribute=48). Retrieved July 2012.
- Henrique Bursztyn and Marius Crainic. Dirac geometry, quasi-Poisson actions and  $D/G$ -valued moment maps. *J. Differential Geom.*, 82(3):501–566, 2009. ISSN 0022-040X. URL <http://projecteuclid.org/getRecord?id=euclid.jdg/1251122545>.
- F. Cantrijn, J. F. Cariñena, M. Crampin, and L. A. Ibort. Reduction of degenerate Lagrangian systems. *J. Geom. Phys.*, 3(3):353–400, 1986. ISSN 0393-0440. doi: 10.1016/0393-0440(86)90014-8. URL [http://dx.doi.org/10.1016/0393-0440\(86\)90014-8](http://dx.doi.org/10.1016/0393-0440(86)90014-8).
- J. F. Cariñena, J. Gomis, L. A. Ibort, and N. Román. Canonical transformations theory for presymplectic systems. *J. Math. Phys.*, 26(8):1961–1969, 1985. ISSN 0022-2488. doi: 10.1063/1.526864. URL <http://dx.doi.org/10.1063/1.526864>.
- José F. Cariñena. Theory of singular Lagrangians. *Fortschr. Phys.*, 38(9):641–679, 1990. ISSN 0015-8208. doi: 10.1002/prop.2190380902. URL <http://dx.doi.org/10.1002/prop.2190380902>.
- José F. Cariñena and Manuel F. Rañada. Blow-up regularization of singular Lagrangians. *J. Math. Phys.*, 25(8):2430–2435, 1984. ISSN 0022-2488. doi: 10.1063/1.526450. URL <http://dx.doi.org/10.1063/1.526450>.
- José F. Cariñena and Manuel F. Rañada. Lagrangian systems with constraints: a geometric approach to the method of Lagrange multipliers. *J. Phys. A*, 26(6):1335–1351, 1993. ISSN 0305-4470. URL <http://stacks.iop.org/0305-4470/26/1335>.
- José F. Cariñena and Manuel F. Rañada. Comments on the presymplectic formalism and the theory of regular Lagrangians with constraints. *J. Phys. A*, 28(3):L91–L97, 1995. ISSN 0305-4470. URL <http://stacks.iop.org/0305-4470/28/L91>.
- José F. Cariñena, Carlos López, and Narciso Román-Roy. Origin of the Lagrangian constraints and their relation with the Hamiltonian formulation. *J. Math. Phys.*, 29(5):1143–1149, 1988. ISSN 0022-2488. doi: 10.1063/1.527955. URL <http://dx.doi.org/10.1063/1.527955>.
- Hernán Cendra and María Etchehoury. Desingularization of implicit analytic differential equations. *J. Phys. A*, 39(35):10975–11001, 2006. ISSN 0305-4470. doi: 10.1088/0305-4470/39/35/003. URL <http://dx.doi.org/10.1088/0305-4470/39/35/003>.
- Hernán Cendra, Jerrold E. Marsden, and Tudor S. Ratiu. Geometric mechanics, Lagrangian reduction, and nonholonomic systems. In *Mathematics unlimited—2001 and beyond*, pages 221–273. Springer, Berlin, 2001a.

- Hernán Cendra, Jerrold E. Marsden, and Tudor S. Ratiu. Lagrangian reduction by stages. *Mem. Amer. Math. Soc.*, 152(722):x+108, July 2001b. ISSN 0065-9266.
- Hernán Cendra, Jerrold E. Marsden, Sergey Pekarsky, and Tudor S. Ratiu. Variational principles for Lie-Poisson and Hamilton-Poincaré equations. *Mosc. Math. J.*, 3(3):833–867, 1197–1198, 2003a. ISSN 1609-3321.
- Hernán Cendra, Jerrold E. Marsden, Sergey Pekarsky, and Tudor S. Ratiu. Variational principles for Lie-Poisson and Hamilton-Poincaré equations. *Mosc. Math. J.*, 3(3):833–867, 1197–1198, 2003b. ISSN 1609-3321. {Dedicated to Vladimir Igorevich Arnold on the occasion of his 65th birthday}.
- Hernán Cendra, Alberto Ibort, Manuel de León, and David Martín de Diego. A generalization of Chetaev’s principle for a class of higher order nonholonomic constraints. *J. Math. Phys.*, 45(7):2785–2801, 2004. ISSN 0022-2488. doi: 10.1063/1.1763245. URL <http://dx.doi.org/10.1063/1.1763245>.
- Hernán Cendra, Jerrold E. Marsden, Tudor S. Ratiu, and Hiroaki Yoshimura. Dirac-Weinstein anchored vector bundle reduction for mechanical systems with symmetry. Preprint, 2011.
- Leon O. Chua and J. D. McPherson. Explicit topological formulation of Lagrangian and Hamiltonian equations for nonlinear networks. *IEEE Trans. Circuits and Systems*, CAS-21:277–286, 1974. ISSN 0098-4094.
- Luis A. Cordero, C. T. J. Dodson, and Manuel de León. *Differential geometry of frame bundles*, volume 47 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1989. ISBN 0-7923-0012-2.
- Jorge Cortés. *Geometric, control and numerical aspects of nonholonomic systems*, volume 1793 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2002. ISBN 3-540-44154-9.
- Jorge Cortés, Manuel de León, Juan C. Marrero, D. Martín de Diego, and Eduardo Martínez. A survey of Lagrangian mechanics and control on Lie algebroids and groupoids. *Int. J. Geom. Methods Mod. Phys.*, 3(3):509–558, 2006. ISSN 0219-8878. doi: 10.1142/S0219887806001211. URL <http://dx.doi.org/10.1142/S0219887806001211>.
- Jorge Cortés, Manuel de León, Juan Carlos Marrero, and Eduardo Martínez. Non-holonomic Lagrangian systems on Lie algebroids. *Discrete Contin. Dyn. Syst.*, 24(2):213–271, 2009. ISSN 1078-0947. doi: 10.3934/dcds.2009.24.213. URL <http://dx.doi.org/10.3934/dcds.2009.24.213>.
- Ted Courant and Alan Weinstein. Beyond Poisson structures. In *Action hamiltoniennes de groupes. Troisième théorème de Lie (Lyon, 1986)*, volume 27 of *Travaux en Cours*, pages 39–49. Hermann, Paris, 1988.

- Theodore James Courant. Dirac manifolds. *Trans. Amer. Math. Soc.*, 319(2):631–661, 1990. ISSN 0002-9947. doi: 10.2307/2001258. URL <http://dx.doi.org/10.2307/2001258>.
- M. Crampin and T. Mestdag. The Cartan form for constrained Lagrangian systems and the nonholonomic Noether theorem. *Int. J. Geom. Methods Mod. Phys.*, 8(4):897–923, 2011a. ISSN 0219-8878. doi: 10.1142/S0219887811005452. URL <http://dx.doi.org/10.1142/S0219887811005452>.
- Mike Crampin and Tom Mestdag. Reduction of invariant constrained systems using anholonomic frames. *J. Geom. Mech.*, 3(1):23–40, 2011b. ISSN 1941-4889. doi: 10.3934/jgm.2011.3.23. URL <http://dx.doi.org/10.3934/jgm.2011.3.23>.
- Manuel de León and David M. de Diego. On the geometry of non-holonomic Lagrangian systems. *J. Math. Phys.*, 37(7):3389–3414, 1996. ISSN 0022-2488. doi: 10.1063/1.531571. URL <http://dx.doi.org/10.1063/1.531571>.
- Manuel de León and Paulo R. Rodrigues. *Methods of differential geometry in analytical mechanics*, volume 158 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam, 1989. ISBN 0-444-88017-8.
- Manuel de León, David Martin de Diego, and Paulo Pitanga. A new look at degenerate Lagrangian dynamics from the viewpoint of almost-product structures. *J. Phys. A*, 28(17):4951–4971, 1995. ISSN 0305-4470. URL <http://stacks.iop.org/0305-4470/28/4951>.
- M. Delgado-Téllez and A. Ibort. On the geometry and topology of singular optimal control problems and their solutions. *Discrete Contin. Dyn. Syst.*, (suppl.): 223–233, 2003. ISSN 1078-0947. Dynamical systems and differential equations (Wilmington, NC, 2002).
- P. A. M. Dirac. Generalized Hamiltonian dynamics. *Canadian J. Math.*, 2:129–148, 1950. ISSN 0008-414X.
- P. A. M. Dirac. Generalized Hamiltonian dynamics. *Proc. Roy. Soc. London. Ser. A*, 246:326–332, 1958. ISSN 0962-8444.
- Paul A. M. Dirac. *Lectures on quantum mechanics*. Belfer Graduate School of Science Monographs Series. Belfer Graduate School of Science, New York, 1964.
- Mark J. Gotay and James M. Nester. Presymplectic Lagrangian systems. I. The constraint algorithm and the equivalence theorem. *Ann. Inst. H. Poincaré Sect. A (N.S.)*, 30(2):129–142, 1979. ISSN 0020-2339.
- Mark J. Gotay and James M. Nester. Generalized constraint algorithm and special presymplectic manifolds. In *Geometric methods in mathematical physics (Proc. NSF-CBMS Conf., Univ. Lowell, Lowell, Mass., 1979)*, volume 775 of *Lecture Notes in Math.*, pages 78–104. Springer, Berlin, 1980.

- Mark J. Gotay and Jędrzej Śniatycki. On the quantization of presymplectic dynamical systems via coisotropic imbeddings. *Comm. Math. Phys.*, 82(3):377–389, 1981/82. ISSN 0010-3616. URL <http://projecteuclid.org/getRecord?id=euclid.cmp/1103920596>.
- Mark J. Gotay, James M. Nester, and George Hinds. Presymplectic manifolds and the Dirac-Bergmann theory of constraints. *J. Math. Phys.*, 19(11):2388–2399, 1978. ISSN 0022-2488. doi: 10.1063/1.523597. URL <http://dx.doi.org/10.1063/1.523597>.
- J. Grabowski, M. de León, J. C. Marrero, and D. Martín de Diego. Nonholonomic constraints: a new viewpoint. *J. Math. Phys.*, 50(1):013520, 17, 2009. ISSN 0022-2488. doi: 10.1063/1.3049752. URL <http://dx.doi.org/10.1063/1.3049752>.
- Xavier Gràcia and Josep M. Pons. Constrained systems: a unified geometric approach. *Internat. J. Theoret. Phys.*, 30(4):511–516, 1991. ISSN 0020-7748. doi: 10.1007/BF00672895. URL <http://dx.doi.org/10.1007/BF00672895>.
- Xavier Gràcia and Josep M. Pons. A generalized geometric framework for constrained systems. *Differential Geom. Appl.*, 2(3):223–247, 1992. ISSN 0926-2245. doi: 10.1016/0926-2245(92)90012-C. URL [http://dx.doi.org/10.1016/0926-2245\(92\)90012-C](http://dx.doi.org/10.1016/0926-2245(92)90012-C).
- Marc Henneaux and Claudio Teitelboim. *Quantization of gauge systems*. Princeton University Press, Princeton, NJ, 1992. ISBN 0-691-08775-X; 0-691-03769-8.
- Darryl D. Holm. *Geometric mechanics. Part I*. Imperial College Press, London, second edition, 2011a. ISBN 978-1-84816-775-9; 1-84816-775-X. Dynamics and symmetry.
- Darryl D. Holm. *Geometric mechanics. Part II. Rotating, translating and rolling*. Imperial College Press, London, second edition, 2011b. ISBN 978-1-84816-778-0; 1-84816-778-4.
- A. Ibort, M. de León, G. Marmo, and D. Martín de Diego. Non-holonomic constrained systems as implicit differential equations. *Rend. Sem. Mat. Univ. Politec. Torino*, 54(3):295–317, 1996. ISSN 0373-1243. Geometrical structures for physical theories, I (Vietri, 1996).
- Alberto Ibort, Manuel de León, Juan C. Marrero, and David Martín de Diego. Dirac brackets in constrained dynamics. *Fortschr. Phys.*, 47(5):459–492, 1999. ISSN 0015-8208. doi: 10.1002/(SICI)1521-3978(199906)47:5<459::AID-PROP459>3.0.CO;2-E. URL [http://dx.doi.org/10.1002/\(SICI\)1521-3978\(199906\)47:5<459::AID-PROP459>3.0.CO;2-E](http://dx.doi.org/10.1002/(SICI)1521-3978(199906)47:5<459::AID-PROP459>3.0.CO;2-E).
- David Iglesias, Juan Carlos Marrero, David Martín de Diego, Eduardo Martínez, and Edith Padrón. Reduction of symplectic Lie algebroids by a Lie subalgebroid and a symmetry Lie group. *SIGMA Symmetry Integrability Geom. Methods Appl.*,

- 3:Paper 049, 28, 2007. ISSN 1815-0659. doi: 10.3842/SIGMA.2007.049. URL <http://dx.doi.org/10.3842/SIGMA.2007.049>.
- Olga Krupková. A geometric setting for higher-order Dirac-Bergmann theory of constraints. *J. Math. Phys.*, 35(12):6557–6576, 1994. ISSN 0022-2488. doi: 10.1063/1.530691. URL <http://dx.doi.org/10.1063/1.530691>.
- Olga Krupková. A new look at Dirac’s theory of constrained systems. In *Gravity, particles and space-time*, pages 507–517. World Sci. Publ., River Edge, NJ, 1996.
- G. H. Livens. On Hamilton’s principle and the modified function in analytical dynamics. *Proc. Roy. Soc. Edinburgh*, (39):113, 1919.
- Giuseppe Marmo, Giovanna Mendella, and Włodzimierz M. Tulczyjew. Constrained Hamiltonian systems as implicit differential equations. *J. Phys. A*, 30(1):277–293, 1997. ISSN 0305-4470. doi: 10.1088/0305-4470/30/1/020. URL <http://dx.doi.org/10.1088/0305-4470/30/1/020>.
- Jerrold E. Marsden and Tudor S. Ratiu. *Introduction to mechanics and symmetry*, volume 17 of *Texts in Applied Mathematics*. Springer-Verlag, New York, 1994. ISBN 0-387-97275-7; 0-387-94347-1.
- Eduardo Martínez. Lie algebroids in classical mechanics and optimal control. *SIGMA Symmetry Integrability Geom. Methods Appl.*, 3:Paper 050, 17 pp. (electronic), 2007. ISSN 1815-0659. doi: 10.3842/SIGMA.2007.050. URL <http://dx.doi.org/10.3842/SIGMA.2007.050>.
- Eduardo Martínez. Variational calculus on Lie algebroids. *ESAIM Control Optim. Calc. Var.*, 14(2):356–380, 2008. ISSN 1292-8119. doi: 10.1051/cocv:2007056. URL <http://dx.doi.org/10.1051/cocv:2007056>.
- B. M. Maschke and A. J. van der Schaft. Note on the dynamics of lc circuits with elements in excess. Memorandum 1426, Faculty of Applied Mathematics of the University of Twente, January 1998.
- B. M. Maschke, A. J. van der Schaft, and P. C. Breedveld. An intrinsic Hamiltonian formulation of network dynamics: nonstandard Poisson structures and gyrators. *J. Franklin Inst.*, 329(5):923–966, 1992. ISSN 0016-0032. doi: 10.1016/S0016-0032(92)90049-M. URL [http://dx.doi.org/10.1016/S0016-0032\(92\)90049-M](http://dx.doi.org/10.1016/S0016-0032(92)90049-M).
- B. M. Maschke, A. J. van der Schaft, and P. C. Breedveld. An intrinsic Hamiltonian formulation of the dynamics of LC-circuits. *IEEE Trans. Circuits Systems I Fund. Theory Appl.*, 42(2):73–82, 1995. ISSN 1057-7122. doi: 10.1109/81.372847. URL <http://dx.doi.org/10.1109/81.372847>.
- Giovanna Mendella, Giuseppe Marmo, and Włodzimierz M. Tulczyjew. Integrability of implicit differential equations. *J. Phys. A*, 28(1):149–163, 1995. ISSN 0305-4470. URL <http://stacks.iop.org/0305-4470/28/149>.



- Luc Moreau and Dirk Aeyels. A novel variational method for deriving Lagrangian and Hamiltonian models of inductor-capacitor circuits. *SIAM Rev.*, 46(1):59–84 (electronic), 2004. ISSN 0036-1445. doi: 10.1137/S0036144502409020. URL <http://dx.doi.org/10.1137/S0036144502409020>.
- N. Mukunda. Time-dependent constraints in classical dynamics. *Phys. Scripta*, 21(6):801–804, 1980. ISSN 0031-8949.
- N. Mukunda. The life and work of P. A. M. Dirac. In *Recent developments in theoretical physics (Kottayam, 1986)*, pages 260–282. World Sci. Publishing, Singapore, 1987.
- Yuri I. Neĭmark and Nikolai A. Fufaev. *Dynamics of Nonholonomic Systems*. Translations of Mathematical Monographs, Vol. 33. American Mathematical Society, Providence, R.I., 1972.
- F. Leon Pritchard. On implicit systems of differential equations. *J. Differential Equations*, 194(2):328–363, 2003. ISSN 0022-0396. doi: 10.1016/S0022-0396(03)00191-8. URL [http://dx.doi.org/10.1016/S0022-0396\(03\)00191-8](http://dx.doi.org/10.1016/S0022-0396(03)00191-8).
- Patrick J. Rabier and Werner C. Rheinboldt. A geometric treatment of implicit differential-algebraic equations. *J. Differential Equations*, 109(1):110–146, 1994. ISSN 0022-0396. doi: 10.1006/jdeq.1994.1046. URL <http://dx.doi.org/10.1006/jdeq.1994.1046>.
- Ray Skinner. First-order equations of motion for classical mechanics. *J. Math. Phys.*, 24(11):2581–2588, 1983. ISSN 0022-2488. doi: 10.1063/1.525653. URL <http://dx.doi.org/10.1063/1.525653>.
- Ray Skinner and Raymond Rusk. Generalized Hamiltonian dynamics. I. Formulation on  $T^*Q \oplus TQ$ . *J. Math. Phys.*, 24(11):2589–2594, 1983a. ISSN 0022-2488. doi: 10.1063/1.525654. URL <http://dx.doi.org/10.1063/1.525654>.
- Ray Skinner and Raymond Rusk. Generalized Hamiltonian dynamics. II. Gauge transformations. *J. Math. Phys.*, 24(11):2595–2601, 1983b. ISSN 0022-2488. doi: 10.1063/1.525655. URL <http://dx.doi.org/10.1063/1.525655>.
- S. Smale. On the mathematical foundations of electrical circuit theory. *J. Differential Geometry*, 7:193–210, 1972. ISSN 0022-040X.
- Jędrzej Śniatycki. Dirac brackets in geometric dynamics. *Ann. Inst. H. Poincaré Sect. A (N.S.)*, 20:365–372, 1974.
- E. C. G. Sudarshan and N. Mukunda. *Classical dynamics: a modern perspective*. Wiley-Interscience [John Wiley & Sons], New York, 1974.
- Héctor J. Sussmann. Orbits of families of vector fields and integrability of distributions. *Trans. Amer. Math. Soc.*, 180:171–188, 1973. ISSN 0002-9947.

- A. J. van der Schaft. Equations of motion for Hamiltonian systems with constraints. *J. Phys. A*, 20(11):3271–3277, 1987. ISSN 0305-4470. URL <http://stacks.iop.org/0305-4470/20/3271>.
- A. J. van der Schaft. Implicit Hamiltonian systems with symmetry. *Rep. Math. Phys.*, 41(2):203–221, 1998. ISSN 0034-4877. doi: 10.1016/S0034-4877(98)80176-6. URL [http://dx.doi.org/10.1016/S0034-4877\(98\)80176-6](http://dx.doi.org/10.1016/S0034-4877(98)80176-6).
- A. J. van der Schaft and B. M. Maschke. On the Hamiltonian formulation of non-holonomic mechanical systems. *Rep. Math. Phys.*, 34(2):225–233, 1994. ISSN 0034-4877. doi: 10.1016/0034-4877(94)90038-8. URL [http://dx.doi.org/10.1016/0034-4877\(94\)90038-8](http://dx.doi.org/10.1016/0034-4877(94)90038-8).
- Alan Weinstein. The local structure of Poisson manifolds. *J. Differential Geom.*, 18(3):523–557, 1983. ISSN 0022-040X. URL <http://projecteuclid.org/getRecord?id=euclid.jdg/1214437787>.
- Alan Weinstein. Lagrangian mechanics and groupoids. In *Mechanics day (Waterloo, ON, 1992)*, volume 7 of *Fields Inst. Commun.*, pages 207–231. Amer. Math. Soc., Providence, RI, 1996.
- Hiroaki Yoshimura and Jerrold E. Marsden. Dirac structures in Lagrangian mechanics. I. Implicit Lagrangian systems. *J. Geom. Phys.*, 57(1):133–156, 2006a. ISSN 0393-0440. doi: 10.1016/j.geomphys.2006.02.009. URL <http://dx.doi.org/10.1016/j.geomphys.2006.02.009>.
- Hiroaki Yoshimura and Jerrold E. Marsden. Dirac structures in Lagrangian mechanics. II. Variational structures. *J. Geom. Phys.*, 57(1):209–250, 2006b. ISSN 0393-0440. doi: 10.1016/j.geomphys.2006.02.012. URL <http://dx.doi.org/10.1016/j.geomphys.2006.02.012>.
- Dmitry V. Zenkov, Anthony M. Bloch, and Jerrold E. Marsden. The energy-momentum method for the stability of non-holonomic systems. *Dynam. Stability Systems*, 13(2):123–165, 1998. ISSN 0268-1110.